



University of Richmond  
**UR Scholarship Repository**

---

Geography and the Environment Faculty  
Publications

Geography and the Environment

---

2009

# Surface Mining and Reclamation Effects on Flood Response of Watersheds in the Central Appalachian Plateau Region

Todd R. Lookingbill  
*University of Richmond*, [tlooking@richmond.edu](mailto:tlooking@richmond.edu)

Joseph R. Ferrari

B. McCormick

P. A. Townsend

K. N. Eshleman

Follow this and additional works at: <http://scholarship.richmond.edu/geography-faculty-publications>

 Part of the [Physical and Environmental Geography Commons](#), and the [Terrestrial and Aquatic Ecology Commons](#)

---

## Recommended Citation

Ferrari, J. R., T. R. Lookingbill, B. McCormick, P. A. Townsend, and K. N. Eshleman. "Surface Mining and Reclamation Effects on Flood Response of Watersheds in the Central Appalachian Plateau Region." *Water Resources Research* 45, no. 4 (2009). doi:10.1029/2008WR007109.

This Article is brought to you for free and open access by the Geography and the Environment at UR Scholarship Repository. It has been accepted for inclusion in Geography and the Environment Faculty Publications by an authorized administrator of UR Scholarship Repository. For more information, please contact [scholarshiprepository@richmond.edu](mailto:scholarshiprepository@richmond.edu).

## Surface mining and reclamation effects on flood response of watersheds in the central Appalachian Plateau region

J. R. Ferrari,<sup>1</sup> T. R. Lookingbill,<sup>1</sup> B. McCormick,<sup>1</sup> P. A. Townsend,<sup>2</sup> and K. N. Eshleman<sup>1</sup>

Received 23 April 2008; revised 14 January 2009; accepted 9 February 2009; published 7 April 2009.

[1] Surface mining of coal and subsequent reclamation represent the dominant land use change in the central Appalachian Plateau (CAP) region of the United States. Hydrologic impacts of surface mining have been studied at the plot scale, but effects at broader scales have not been explored adequately. Broad-scale classification of reclaimed sites is difficult because standing vegetation makes them nearly indistinguishable from alternate land uses. We used a land cover data set that accurately maps surface mines for a 187-km<sup>2</sup> watershed within the CAP. These land cover data, as well as plot-level data from within the watershed, are used with HSPF (Hydrologic Simulation Program-Fortran) to estimate changes in flood response as a function of increased mining. Results show that the rate at which flood magnitude increases due to increased mining is linear, with greater rates observed for less frequent return intervals. These findings indicate that mine reclamation leaves the landscape in a condition more similar to urban areas rather than does simple deforestation, and call into question the effectiveness of reclamation in terms of returning mined areas to the hydrological state that existed before mining.

**Citation:** Ferrari, J. R., T. R. Lookingbill, B. McCormick, P. A. Townsend, and K. N. Eshleman (2009), Surface mining and reclamation effects on flood response of watersheds in the central Appalachian Plateau region, *Water Resour. Res.*, 45, W04407, doi:10.1029/2008WR007109.

### 1. Introduction

[2] Watershed hydrology can be altered by changes in land use and land cover (LULC). A number of studies have shown that timber harvest [Hornbeck *et al.*, 1970], urbanization [Hollis, 1975; Burges *et al.*, 1998; Wissmar *et al.*, 2004], changes in agricultural management [Potter, 1991; Allan *et al.*, 1997] and surface mining [Negley and Eshleman, 2006] can shift the hydrologic balance from subsurface to surface flow, yielding changes in rainfall/runoff ratios and increased flood frequency and magnitude. These hydrological responses can be difficult to predict but have important implications for the mitigation of possible flood damage to human life, property, and aquatic and riparian biota [Eshleman, 2004; Wissmar *et al.*, 2004].

[3] Between 1975 and 2000 the most significant changes in LULC in the central Appalachian Plateau (CAP) of the eastern United States were related to surface mining of bituminous coal [Loveland *et al.*, 2003]. In northern and central Appalachia, over  $1.1 \times 10^6$  ha are listed as undergoing active mining operations [Office of Surface Mining, 2004], and mining operations are expected to increase to meet increased energy demands. Surface mining of coal involves removal and storage of the uppermost soil horizon (topsoil), removal of lower soil layers and rock (overburden) to expose

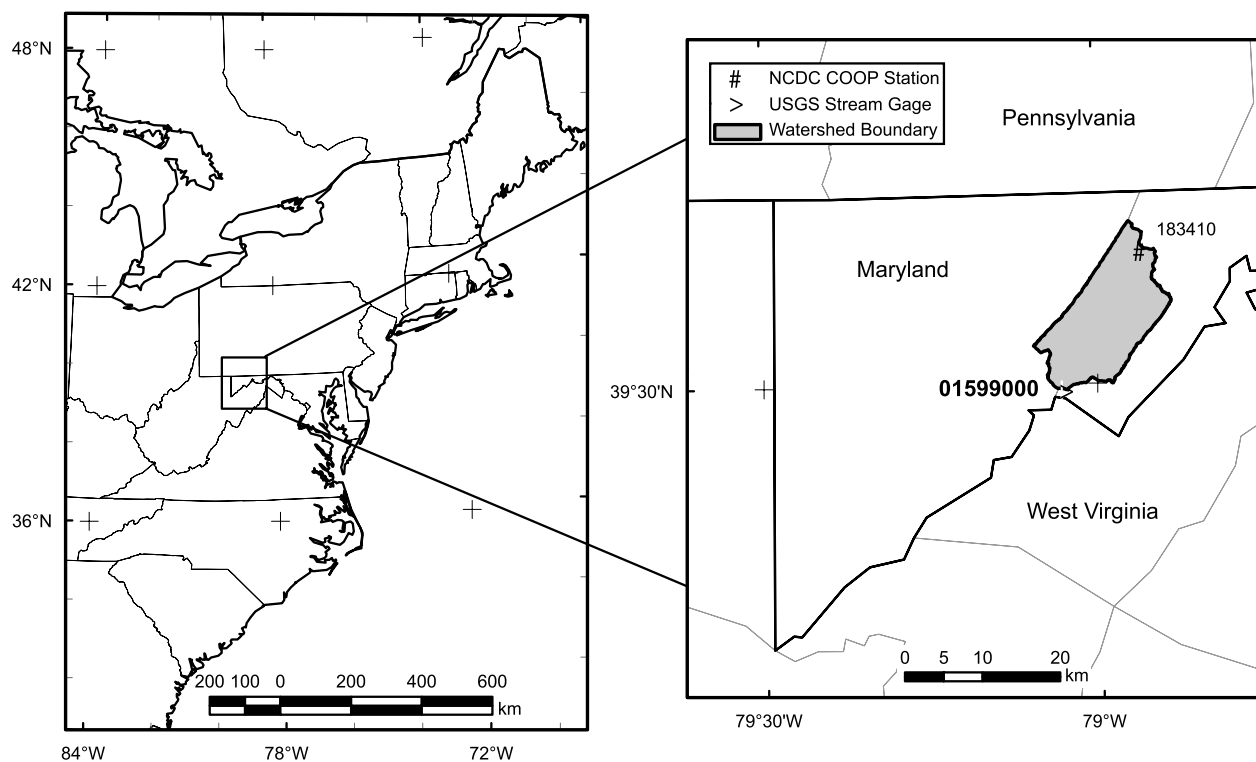
coal seams, followed by removal of coal deposits. The Surface Mine Control and Reclamation Act (SMCRA) of 1977 [U.S. Congress, 1977] requires mine operators to return surface mined sites to the approximate premining contours and to acceptable LULC. This is performed by replacing the overburden and then the topsoil using large earthmovers, followed by vegetation planting or reseeding.

[4] Reclamation using earthmovers substantially compacts soil layers. Compaction increases soil bulk density and decreases porosity and infiltration [Chong *et al.*, 1986]. Infiltration rates on new and reclaimed surface mines can be as much as an order of magnitude smaller than those measured at nearby, undisturbed plots [Jorgensen and Gardner, 1987; Potter *et al.*, 1988; Negley and Eshleman, 2006]. Reduced infiltration alters flow rates into lower soil profiles and groundwater flows [Pederson *et al.*, 1980]. Increased bulk density in surface-mined soils has been shown to yield less than favorable conditions on sites intended for return to a forested condition [Bussler *et al.*, 1984].

[5] The hydrologic impact of surface mining and reclamation has been well documented empirically at the scale of small (0.1–1.0 km<sup>2</sup>) catchments [e.g., Bonta *et al.*, 1997; Negley and Eshleman, 2006], but the effect at broader scales (e.g., 100–1000 km<sup>2</sup>) has not been widely investigated. Hydrologic models can be used to make investigations at broader scales, but modeling approaches have been limited in application due to lack of accurate data on the location and size of reclaimed surface mines. Lumped parameter and spatially explicit hydrologic models require accurate LULC data if they are to be used to determine the effect of mine-related activity. Reclaimed mines may be decades old and

<sup>1</sup>Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, Maryland, USA.

<sup>2</sup>Russell Laboratories, University of Wisconsin-Madison, Madison, Wisconsin, USA.



**Figure 1.** Location of Georges Creek watershed in western Maryland.

covered with vegetation layers, making them nearly indistinguishable from land uses such as agriculture, pasture, shrub/scrub, or forest in mapping projects such as the Multi-Resolution Land Characteristics Consortium National Land Cover Data [Homer *et al.*, 2004].

[6] In this research, we take advantage of a recent highly detailed assessment of LULC (focusing on active and reclaimed surface mines) for a portion of the CAP to assess the effects of surface mining and reclamation on flood response. We calibrate the Hydrologic Simulation Program-Fortran (HSPF) over three 10-year time periods for a 187-km<sup>2</sup> representative watershed from the region, and use the calibrated model to explore the sensitivity of flood response to simulated increases in surface mine activity using 50 years of atmospheric input data. Our objectives were to determine the hydrologic response of watersheds to surface mining activities relative to other forms of LULC conversions and to assess the current state of the focal watershed relative to any potential thresholds of change. We chose to use simulations for this purpose because comparison of flood frequency curves (FFCs) using empirical data premining and postmining within Georges Creek yields equivocal results, whereas event-based analyses indicate higher stormflows as a result of mining (B. C. McCormick *et al.*, Detection of flooding response at the river basin scale enhanced by land use changes, submitted to *Water Resources Research*, 2008). Problems with the FFC approach (using empirical data from different time periods) include differences in rainfall premining and postmining and nonstationary LULC postmining. Our simulation efforts aimed at creating simulated FFCs for various levels of fixed amounts of mine activity, thus removing the effect of nonstationarity, and subject to a single 50-year atmospheric

data set, thus removing the influence of varying environmental inputs.

## 2. Methods

### 2.1. Study Area and Watershed Delineation

[7] The Georges Creek watershed (30°35'N, 79°00'W) is located in western Maryland, on the central Appalachian Plateau of the eastern United States (Figure 1). A 30-m digital elevation model (DEM) was obtained from the U.S. Geological Survey (USGS; <http://seamless.usgs.gov>) and used for watershed delineation, with the pour point set to the USGS gage at Franklin, Maryland (USGS gage 01599000). The U.S. Environmental Protection Agency (EPA) Better Assessment Science Integrating Point and Non-Point Sources (BASINS) 3.1 software [U.S. EPA, 2001] was used to delineate the watershed. Nine subwatersheds were delineated with BASINS for HSPF modeling purposes with a mean subwatershed area of 20 km<sup>2</sup>.

### 2.2. Land Use/Land Cover (LULC)

[8] Land use and land cover change was delineated by Townsend *et al.* [2009] for the study area using Landsat imagery from 1976, 1987, and 1999. Because of significant inaccuracies in mapping mine-related classes (>48% incorrect) in the National Land Cover Data sets (NLCD) of 1992 and 2001 [Vogelmann *et al.*, 2001; Homer *et al.*, 2004], the LULC images were classified using a logical decision tree to assess land cover type based on both spectral characteristics and transitions through time (Table 1). This was necessitated by the fact that without ancillary information, reclaimed grasslands can be indistinguishable from pasture and even some agricultural types. Using temporal information in addition to the spectral data, a pixel that moves from

**Table 1.** Percent Coverage of the Georges Creek Watershed by Land Use/Land Cover Class for 1976, 1987, and 1999

Land Use/Land Cover	1976	1987	1999
Urban	3.4	3.7	4.2
Agriculture	10.0	10.8	9.4
Forest	76.3	70.7	70.0
Active mine	5.4	2.1	1.9
Reclaimed mine	4.9	12.8	14.5
Open water	0.01	0.01	0.01

forest to bare to grassland over three time periods would be correctly classified as forest, active mine, and then reclaimed mine. In addition, the maps used data on mine permits to positively distinguish mined lands from other cover types having similar spectral responses and ambiguous transitions. The accuracy of the *Townsend et al.* [2009] classification for Georges Creek was assessed using aerial photography and was >90% for the 1999 image, with 93.5% producer's accuracy and 100% user's accuracy for distinguishing mine lands. Overall map accuracy for 1976 and 1987 was comparable to 1999, but with slightly lower accuracy for mine classes (85% producer's and user's accuracy in 1976 and 84.4%/96.4% producer/user's accuracy for 1987). These land cover maps are the most accurate available representation of reclaimed mine lands for Georges Creek, and likely any other similarly sized watersheds in the Central Appalachians. Classified images for 1976, 1987, and 1999 were used as LULC input data for hydrologic simulations.

### 2.3. Watershed Modeling Using HSPF

[9] The Hydrologic Simulation Program-Fortran (HSPF) is the core watershed simulation module of the BASINS 3.1 software [Atkins *et al.*, 2005]. We chose HSPF for our work because it has a rich history as a tool for predicting changes in watershed response due to changes in land cover [e.g., *Brun and Band*, 2000; *Coon*, 2003; *Doherty and Johnston*, 2003]. Although a lumped parameter model, HSPF is considered to be moderately physically based [Gallagher and Doherty, 2007]. We used BASINS 3.1 to generate the main input file for HSPF simulations, called the User Control Input (UCI) file. We then modified the UCI file for use with the Parameter Estimation Software, PEST [Doherty, 2005], and the corresponding command line variant of HSPF 12, called XHSPFX.

### 2.4. Streamflow and Atmospheric Data

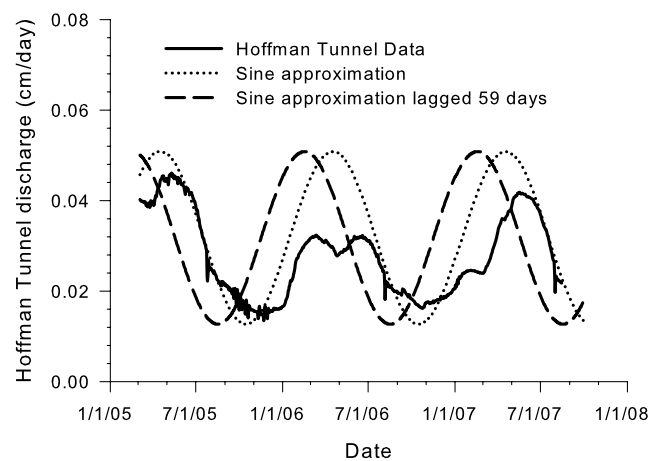
[10] Average daily streamflow data were obtained from the USGS gage station located at Franklin, Maryland (USGS gage 01599000) for the water years 1954 to 2004. Daily records for minimum/maximum air temperatures, precipitation, and snow/ice pack depth were obtained for the 50-year time period from the Frostburg cooperative observer station, located within the Georges Creek watershed. HSPF operates internally on an hourly time step. Because hourly measurements were not recorded at the observer station, we disaggregated daily rainfall to hourly values using hourly precipitation data from the nearby Savage Dam observer station and from our own gage located along Matthews Run, within the Georges Creek watershed. Sixty percent of the daily rainfall events (over the 50-year time span) had

corresponding hourly data to guide disaggregation. The remaining 40% were disaggregated using a default triangular distribution. Disaggregation was performed using the DISAGGREGATE tool in the HSPF weather data management utility WDMUTIL [Hummel *et al.*, 2001].

### 2.5. Additional Considerations for Surface-Mined Watersheds in the CAP

[11] Georges Creek contains complex underground tunnel systems and drainages related to deep shaft mining, a situation common among watersheds in the CAP [e.g., *Atkins et al.*, 2005]. One manifestation of underground mining activities in Georges Creek is the Hoffman Tunnel (HT), a 3.2-km tunnel connected to over 8 km of underground diversion ditches and auxiliary tunnels. The tunnel was completed in 1907 to gravity drain the Pittsburgh coal seam [Maryland Department of Natural Resources, 2001] and acts as a transbasin diversion from the Georges Creek watershed to the adjacent Braddock Run watershed. Recorded annual minimum and maximum flows from the tunnel are in the range of 0.013–0.05 cm/d and have been observed to be relatively stable over time [Maryland Department of Natural Resources, 2001].

[12] We were able to obtain daily flow records for the Hoffman Tunnel for the time period 1 March 2005 to 30 September 2007 D. Welsch, unpublished data, 2008). The daily flow records indicate a seasonal flow pattern with peak flows in April/May and flow minima in October. Further investigation of the Hoffman Tunnel led us to model this drainage as a delayed storage mechanism. Correlogram analysis indicated that Hoffman Tunnel discharge is correlated with magnitude of Georges Creek outflow lagged 59 days ( $r^2 = 0.27$ ). For simulation purposes we assigned a sinusoidal flow to the Hoffman Tunnel discharge with minimum and maximum flows set to their historical values (0.013 and 0.05 cm/d respectively), with peak values lagged 59 days prior to the peaks observed over the 2005–2007 data set (Figure 2). As the Hoffman Tunnel flows have been



**Figure 2.** Hoffman Tunnel discharge. Correlogram analysis indicated Hoffman Tunnel discharge was predictable using Georges Creek flows lagged 59 days ( $r^2 = 0.27$ ). Date range shown is the length for which daily observations were available. For simulation purposes the lagged sine approximation was extended over the 1954–2004 time span for use in 10-year calibrations and 50-year simulations.

**Table 2.** Parameters Used in Calibration of the HSPF Model

Parameter	Description	Units	Range
AGWRC	groundwater recession rate	none	0.85–0.999
CEPSC	interception storage capacity parameter	cm	0.025–1.02
DEEPPFR	fraction of infiltrating water lost to deep aquifers	none	0.0–0.2
INFIL	index to mean soil infiltration rate	cm/h	0.025–1.27
INTFW	interflow inflow rate	none	1.0–10.0
IRC	interflow recession coefficient	none	0.3–0.85
KMELT	parameter used to differentiate rain/snow/ice for degree-day approach to snow calculations	cm/d <sup>-°C</sup>	0.09–0.69
KVARY	groundwater recession flow parameter used for nonlinear groundwater recession	1/cm	0.002–2.0
LZETP	index to lower zone evapotranspiration	none	0.1–0.9
LZSN	lower zone nominal soil moisture storage	cm	2.54–38.1
UZSN	upper zone nominal soil moisture storage	cm	0.13–5.1

considered to be stable in their magnitudes and seasonal patterns, we extended this lagged sinusoidal approximation across the 50-year time span used for simulations.

**2.6. Calibration**

[13] Calibration included the following steps: specification of parameters to vary during calibration and their ranges; specification of goodness of fit criteria for determining when a parameter set meets calibration goals; selection of a parameter set to meet those goals; and quantification of predictive uncertainty that results from parameter uncertainty.

**2.6.1. Parameter Selection**

[14] We selected 11 parameters (Table 2) as most important for the calibration of hydrology in mined watersheds [Lumb et al., 1994; U.S. EPA, 2000; Atkins et al., 2005]. These parameters influence partitioning of water flows, such as overland, interflow, and ground flow, within the HSPF model (Figure 3). Full information on these parameters and their use in the HSPF model is described in HSPF publications [Bicknell et al., 2001; U.S. EPA, 2000]. All

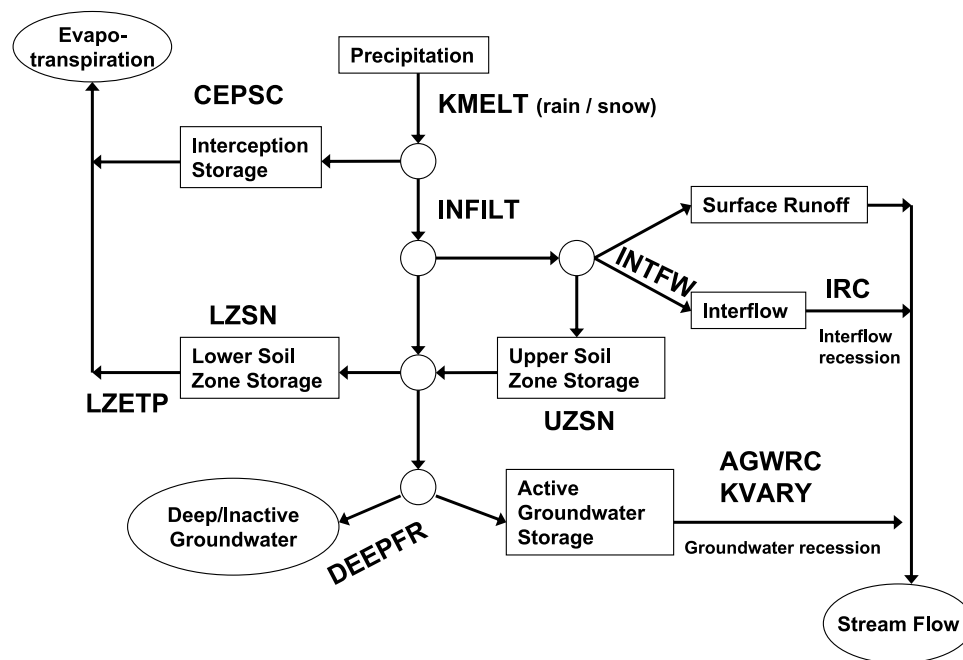
other parameters were assigned values as suggested by U.S. EPA [2000]. Monthly values were assigned to the parameters CEPSC and LZETP to reflect seasonal variations in interception storage and evapotranspiration, respectively. During calibration, the groundwater (AGWRC) and interflow (IRC) recession parameters were not used directly. Instead, the following transformed parameters (equations (1) and (2)) were varied during calibration to offset possible numerical instability when using the native parameters, as suggested by Doherty and Johnston [2003].

$$AGWRC^* = AGWRC / (1.0 - AGWRC) \tag{1}$$

$$IRC^* = IRC / (1.0 - IRC) \tag{2}$$

**2.6.2. Parameter Ties**

[15] Each of our six land cover classes (Table 1) has its own unique parameters in the HSPF model. Calibration can approach multiclass parameterization (11 parameters by six



**Figure 3.** Conceptual diagram of flow routing across pervious land segments in the Hydrologic Simulation Program-Fortran (HSPF), adapted from Atkins et al. [2005]. Text outside of boxes/ovals indicates parameters. Parameters INFIL, CEPSC, and LZETP varied specifically for mined lands.

classes) through parameter ties. Parameter ties allow users to calibrate one set of LULC parameters (e.g., forest, the dominant land use in Georges Creek) and then assign parameter values to corresponding LULC parameter pairs based on knowledge of their expected relative differences. Parameter ranges for forested lands (Table 2) were assigned the minimum and maximum values recommended by *U.S. EPA* [2000]. Ties among forest parameters and those of other cover types were assigned based on published data when available, including suggestions by *U.S. EPA* [2000]; otherwise we relied on our collective experience working in these systems. For example, data for nested subwatersheds in the Georges Creek basin indicate that reclaimed surface mines have infiltration rates of the order of 3% of undisturbed forests [Negley and Eshleman, 2006]. Interception storage (CEPSC) and evapotranspiration (LZETP) ties were assigned to each nonforest land use based on their vegetative cover.

### 2.6.3. Definition of the Objective Function

[16] The goodness-of-fit-criteria used to assess performance of calibrations was based on an objective function (equation (3)) that exhibits smaller values the closer the simulated flows are to observed values:

$$\phi = \sum_{i=1}^3 w_i \phi_i, \quad (3)$$

where  $w_i$  = weight applied to the  $i$ th component  $\phi_i$ . The three components of the objective function are as follows:

$$\phi_1 = \sqrt{\sum_{j=1}^N (q'_j - q_j)^2}, \quad (4)$$

where  $q'_j$  and  $q_j$  are the HSPF predicted and the observed outflows (cm/d), respectively, and  $N$  is the total number of days in the calibration time period;

$$\phi_2 = \sqrt{\sum_{j=1}^M (V'_j - V_j)^2}, \quad (5)$$

where  $V'_j$  and  $V_j$  are the HSPF predicted and the observed flow volumes (cm), respectively, calculated over  $M$  flood dates. Flood dates were chosen as three to five of the largest runoff events for each water year in the calibration time period. Dates for each event included one or more days prior to each peak flow to several days following each peak.

[17] To account for the effect of snow/ice pack on simulated runoff, we assigned the third component of the objective function as

$$\phi_3 = \sqrt{\sum_{j=1}^N (p'_j - p_j)^2}, \quad (6)$$

where  $p'_j$  and  $p_j$  are the HSPF predicted and the observed snow/ice pack depth (cm).

[18] The first part of the objective function has similarities to the Nash-Sutcliffe efficiency, in that the squared error terms yield substantially larger contributions for peak flow errors versus smaller contributions from differences in

low flow events [Criss and Winston, 2008]. This is advantageous as our study is aimed at understanding peak flow events. However, performing a calibration to minimize this form of objective function is not always guaranteed to provide optimal peak flow matching [Gallagher and Doherty, 2007]. Therefore the objective function components  $\phi_2$  and  $\phi_3$  were added to provide greater emphasis on stormflows and the effect of snow/ice pack on storm runoff. Weights of  $w_1 = 1.0$ ,  $w_2 = 2.0$ , and  $w_3 = 0.025$  were assigned for use in equation (3) to ensure all parts of the objective function were of the same order of magnitude during calibration, with the stormflow component  $\phi_2$  approximately twice the value of the other components. The goal of subsequent calibration was to minimize the objective function  $\phi$  and to quantify uncertainty associated with parameter nonuniqueness.

[19] The first part of the calibration used Monte Carlo (MC) simulations in which 5000 randomly selected parameter sets were drawn assuming uniform parameter distributions within their ranges (Table 2). Each parameter set was used to run a simulation across the calibration time period and the corresponding value of  $\phi$  calculated. The resultant mapping of the solution space was equivalent to that of the Generalized Likelihood Uncertainty Estimation method (GLUE [Beven and Binley, 1992]), which has been used in other studies to calibrate HSPF [e.g., Ewen et al., 2006; Jia and Culver, 2008].

[20] The parameter estimation software PEST [Doherty, 2005], and its accompanying driver PD\_MS2 [Doherty, 2007], were designed to use the results of MC simulations to guide application of the GML, or Gauss-Marquardt-Levenberg technique [Levenburg, 1944; Marquardt, 1963]. The GML algorithm takes an initial parameter set and numerically optimizes parameter values to obtain a  $\phi$  value smaller than that corresponding to the initial parameter set. The goal of calibration is to determine the parameter set corresponding to the global minima value of  $\phi$ . The GML technique is very efficient at finding minima in  $\phi$  space, but it is also prone to convergence on local minima (as opposed to the global minima) if the initial parameter values yield  $\phi$  close to a local minima. Therefore the PD\_MS2 (PEST) driver first selects the parameter set that yielded the smallest  $\phi$  from the 5000 MC runs to initiate the GML search. Multiple additional parameter sets from among the 5000 MC simulations were also used to initiate different GML searches. These sets were selected based on Euclidean distance (in parameter space) relative to the parameter set associated with the smallest  $\phi$  used for the initial GML search. The reader is referred to Doherty and Johnston [2003] and Gallagher and Doherty [2007] for full details on the global minima search methodology used by PD\_MS2. Each selected parameter set was used to initiate a GML optimization, and the search was stopped when 50 successive searches did not yield a  $\phi$  reduction of 3% or greater, with the smallest value of  $\phi$  equated to the global minima.

[21] Complex models with more than a few parameters rarely have a bowl-shaped solution space for  $\phi$  with a well-defined “pit” corresponding to the global minima. Minima typically lie along a “valley” [Gallagher and Doherty, 2007]. Along this valley may be multiple parameter sets that yield similar  $\phi$  and can be said to equally calibrate the model. Therefore, of the optimized parameter sets identified during the search for the global minima, we retained that

**Table 3.** Calibration Statistics for Georges Creek Calibration (Water Years 1994–2004) Using the DEEPFR Parameter to Approximate Hoffman Tunnel Discharge (Calibration 1) and With DEEPFR Held at 0.0, With Explicit Handling of the Hoffman Tunnel (Calibration 2)<sup>a</sup>

Statistic	Calibration 1	Calibration 2
Nash-Sutcliffe efficiency	0.73 (0.71–0.74)	0.75 (0.74–0.76)
Overall water balance error (%)	25.6 (18.3–32.9)	3.4 (–1.3 to 8.1)
Mean yearly water balance error (%)	25.6 (22.0–27.8)	2.8 (–2.1 to 7.3)
Mean monthly water balance error (%)	71.8 (59.7–83.9)	6.6 (0.7–12.4)
Lowest 50% flow error (%)	74.8 (62.2–86.0)	–10.8 (–16.5 to –5.7)
Highest 10% flow error (%)	7.31 (2.8–11.6)	5.8 (3.0–9.0)
Mean storm volume error (%)	1.7 (–4.0 to 7.4)	–3.3 (–7.1 to 0.5)
Mean storm peak flow error (%)	–15.4 (–18.8 to –12.0)	–14.5 (–16.9 to –12.0)

<sup>a</sup>Values are for statistics using  $\bar{q}_i$  time series (equation (7)); entries in parentheses indicate predictive uncertainty intervals.

associated with the smallest (optimized) value of  $\phi$ , as well as any others that had  $\phi$  values within 2% of the minima, for a total of  $k$  parameter sets.

[22] The predicted daily flows from a calibration were then calculated as the average of flow values for that day across  $k$  simulations each driven by its associated parameter set,

$$\bar{q}_i = \frac{\sum_{j=1}^k (q_i)_j}{k}, \quad (7)$$

where the index  $i$  refers to the date (day) during the calibration time period and the index  $j$  refers to the parameter set used by HSPF to calculate the individual  $q_i$  time series. The standard deviation (*Stdev*) and 95% confidence intervals were also calculated for each daily flow across  $k$  parameterizations. The upper and lower intervals,  $(q_i^* = \bar{q}_i \pm 95\%$  confidence interval) represent the predictive uncertainty of the parameterization based on nonuniqueness of  $k$  parameter sets with  $\phi$  values within 2% of the global minima found during calibration. These intervals represent “conditional uncertainties” relative to the criteria set for parameter acceptance [Beven, 2006].

[23] For each calibration, additional statistics were calculated using the associated  $\bar{q}_i$  and  $q_i^*$  time series. These statistics included Nash-Sutcliffe efficiency [Krause et al., 2005] and percent differences (between simulated and observed) in overall water balance, yearly water balance, monthly water balance, lowest 50% of flows, highest 10% of flows, overall flood volume, and peak flood flow (resolved to the daily basis for which streamflow data were available). Flood statistics were calculated for the  $M$  peak flow events for each calibration time period and represent flood response under a wide range of rainfall intensities and antecedent soil moisture conditions.

#### 2.6.4. Calibration Date Ranges

[24] We calibrated the HSPF model for 10 water years (1994–2004) centered on 1999 LULC. Two variations of this calibration were performed. The first used DEEPFR as a variable to account for deep losses. The second deactivated the DEEPFR parameter and instead compared the simulated flow to the sum of the observed flow and the Hoffman Tunnel (HT) approximation. The purpose was to assess whether our HT approximation of deep losses yielded better performance than the HSPF mechanism driven by DEEPFR. The method chosen for handling deep losses was then used to calibrate HSPF across water years 1971–1981

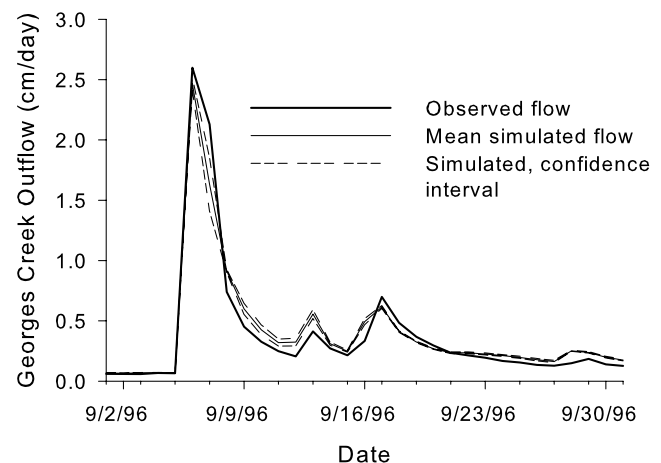
and 1982–1992, using LULC data from image dates 1976 and 1987, respectively. Each of the calibration simulations used a 1-year spin-up time for HSPF. Ten-year spans were selected to incorporate wide variety in atmospheric driving conditions for hydrologic response, specifically for peak flow events. Calibration time spans were centered on LULC image dates to minimize effects of nonstationarity in land use as suggested by Gutierrez-Magness [2005]. The calibration time periods centered on 1976, 1987, and 1999 had  $M = 51, 35,$  and  $44$  storm events identified, respectively, for use in calibration (equation (5)) and for calculation of stormflow statistics.

#### 2.6.5. Model Evaluation

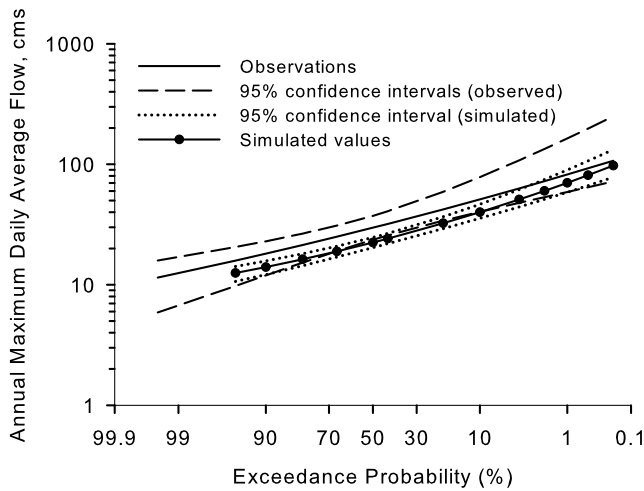
[25] If parameters obtained for a given time period are robust, they should yield satisfactory simulations for other time periods [Jacomino and Fields, 1997]. To evaluate the ability of the model to satisfactorily simulate flows across other time periods, we performed the following experiment. Parameter sets from each calibration (1971–1981, 1982–1992, and 1994–2004) were used to drive the simulations for each of the three time periods. The Nash-Sutcliffe efficiency, mean storm mass balance, and mean storm peak daily flow errors were calculated for each simulation.

#### 2.7. Land Use/Land Cover (LULC) Change Simulation

[26] Our goal was to assess flood response as a function of  $p$ , the percent of the watershed covered by reclaimed mine. The experience in many parts of the CAP is a long-



**Figure 4.** Representative plot of simulated and observed flows for Georges Creek for the month of September 1996, including predictive uncertainty intervals.



**Figure 5.** Log Pearson III flood frequency curve for Georges Creek, simulated data (by equation (7)) and observed, for the time span 1994–2004. Largest probabilities for simulated data were not plotted, as the values fell below the cutoff threshold of the program PEAKFQ.

term presence of minor amounts of active mines which are subsequently reclaimed. The result is that while active mines may only cover a small proportion of a watershed at any given point in time, the legacy of years of mining can affect much greater proportions of the watershed as reclaimed mine lands accumulate over time.

[27] Increased mining was simulated by converting forest to reclaimed mine while keeping the percent of the watershed covered by other LULC (e.g., active mine, agriculture, urban, water) constant at values for the 1999 image date. In 1999, 70% of the watershed was forested (Table 1). We sequentially converted forest to reclaimed mine in increments of 7% until all forest was converted to reclaimed mine. Two more scenarios were also tested, the first reducing reclaimed mine coverage to 7.5% (from 14.5% in 1999) and the second with reclaimed mine coverage set to 0%, with corresponding increases in forest coverage. Our simulations thus included a span of reclaimed mine coverage from  $p = 0\%$  to  $p = 84.5\%$ . Within the HSPF modeling framework, these shifts in LULC were accomplished by varying the amounts of forest and reclaimed mine in the SCHEMATIC block of the UCI file. Class conversion was split evenly among each of the nine subwatersheds defined during watershed delineation.

[28] Simulations for each LULC change scenario were performed using 50 years (1954–2004) of climate conditions. Data recorded from the simulations included daily

flow data across parameterizations ( $\bar{q}_i$ ) and predictive uncertainty intervals ( $q_i^*$ ). Log Pearson III flood frequency distributions were calculated for annual maximum daily modeled flows using PEAKFQ [Flynn et al., 2006] per Interagency Advisory Committee on Water Data [1982] guidelines. The flood frequency distributions were then used to extract the magnitude of 2-, 10-, and 50-year flood events and plotted against the percent  $p$  of Georges Creek affected by mining and subsequent reclamation.

### 3. Results

#### 3.1. Deep Losses/Hoffman Tunnel

[29] Calibration of Georges Creek watershed using 1999 land use/land cover (LULC) data yielded similar Nash-Sutcliffe efficiency and peak daily stormflow errors using both the DEEPFR parameter and our explicit representation of the Hoffman Tunnel (Table 3). However, the remaining mass balance errors using the DEEPFR parameter were between 2 and 10 times the values obtained using the explicit HT representation. We retained explicit handling of the HT for further calibrations based on the overall goodness of fit relative to use of the DEEPFR parameter. Figure 4 provides an illustration of the simulated  $\bar{q}_i$  time series with predictive uncertainty intervals using explicit handling of the HT.

#### 3.2. Model Evaluation for Fixed LULC

[30] The simulated data for the 10-year period (1994–2004) were compared with observations using log Pearson III flood frequency curves (Figure 5). Simulated data fell within the confidence limits of the observed data, and the trends in observed and simulated data were similar. Differences are in part attributable to comparing a fixed land use (simulations with LULC set at 1999 levels; Table 1) to empirical data over a period with nonstationary LULC.

#### 3.3. Model Evaluation Across Land Use Change

[31] We calibrated the HSPF model for the 10-year periods centered on image dates 1976, 1987, and 1999 using our sine wave approximation to HT discharge. The optimized objective function value (by equation (3)) was approximately 20% below the smallest value determined by Monte Carlo simulation for each of the three calibrations. For each calibration we retained between  $k = 15$  and  $k = 20$  parameter sets. Parameter sets for each calibration were used to generate the  $\bar{q}_i$  time series as per equation (7) and predictive uncertainty intervals across each of the three calibration time periods. Simulations run using the parameterization for 1994–2004 yielded diagnostic statistics (Tables 4, 5, and 6) for the other two time periods (1971–1981 and 1982–1992) comparable to values attained using parameter sets derived explicitly for those time periods. As each of the time periods had different

**Table 4.** Nash-Sutcliffe Efficiency for Simulations Using Parameters Calibrated for Each Simulation Time Period<sup>a</sup>

Nash Sutcliffe Efficiency, 10-Year Simulation Period	Using Parameters From Calibration Time Period		
	1971–1981	1982–1992	1994–2004
1971–1981	<b>0.72</b> (0.71–0.72)	<b>0.67</b> (0.66–0.67)	<b>0.72</b> (0.71–0.72)
1982–1992	<b>0.69</b> (0.68–0.70)	<b>0.77</b> (0.76–0.78)	<b>0.77</b> (0.75–0.78)
1994–2004	<b>0.69</b> (0.68–0.69)	<b>0.75</b> (0.74–0.75)	<b>0.75</b> (0.74–0.76)

<sup>a</sup>Bold values are for statistics using  $\bar{q}_i$  time series (equation (7)); values in parentheses indicate predictive uncertainty intervals.



**Table 5.** Mean Storm Volume Errors for Simulations Using Parameters Calibrated for Each Simulation Time Period<sup>a</sup>

Mean Storm Volume Errors (%), 10-Year Simulation Period	Using Parameters From Calibration Time Period		
	1971–1981	1982–1992	1994–2004
1971–1981	<b>-5.4</b> (-7.5 to -3.2)	-7.1 (-11.1 to -3.2)	<b>-8.3</b> (-11.6 to -5.1)
1982–1992	<b>7.0</b> (5.2–8.8)	-0.1 (-2.6 to 2.8)	<b>5.2</b> (1.6–8.9)
1994–2004	<b>-1.9</b> (-4.0 to 0.2)	<b>-6.1</b> (-8.9 to -3.3)	<b>-3.3</b> (-7.1 to 0.48)

<sup>a</sup>Bold values are for statistics using  $\bar{q}_t$  time series (equation (7)); values in parentheses indicate predictive uncertainty intervals.

LULC, these results provided confidence that the parameterization for 1994–2004 was robust across land use change.

#### 3.4. Land Use/Land Cover (LULC) Change Simulation

[32] The log Pearson III flood frequency plots were created using 50 years of simulated streamflow (representative plots shown in Figure 6) for each LULC change scenario, ranging from  $p = 0\%$ , and from  $p = 7.5\%$  to  $p = 84.5\%$  in increments of 7%. In all LULC experiments the amounts of all other classes were kept constant at 1999 values (Table 1). Magnitude of daily peak flow at 2-, 10-, and 50-year return intervals was plotted against  $p$ . Results (Figure 7) indicate a somewhat linear response in peak flow at the stated return intervals. Linear regressions were fitted to the 2-, 10-, and 50-year plots and the slopes of the regressions calculated as 0.41, 0.80, and 1.26, respectively ( $r^2 > 0.99$  for all three regressions). These values represent the rate of change of 2-, 10-, and 50-year peak average daily flows as a function of  $p$ .

#### 4. Discussion

[33] Surface mining and subsequent reclamation is the dominant vector of land use/land cover change (LULC) in the central Appalachian Plateau (CAP) [Loveland *et al.*, 2003]. Although many studies have been done on the flood response of mined sites at the plot scale (0.1–1.0 km<sup>2</sup>), few have been done at the scale of larger watersheds that contribute to major river systems (100–1000 km<sup>2</sup>). Our goal was to investigate the flood response of the 187-km<sup>2</sup> Georges Creek watershed as a function of increased surface mining and subsequent reclamation. Flood response was evaluated at the daily time step due to unavailability of hourly data for this and many other watersheds in the CAP. Our results do not provide data on instantaneous flood peaks, but trends in exceedance probabilities at hourly time steps are often more pronounced than when resolved to daily time steps [Samuel and Sivapalan, 2008]. Therefore our results are conservative and may underestimate trends expected for instantaneous peak flood response expected from surface mined watersheds such as Georges Creek.

[34] Georges Creek was chosen because at this scale of analysis, this watershed has one of the largest proportions of reclamation of any in the CAP. Empirical studies within this watershed have shown that surface-mined sites experienced a tripling of total storm runoff and doubled peak hourly flows relative to adjacent, forested sites [Negley and Eshleman, 2006]. More recent empirical work by B. C. McCormick *et al.* (submitted manuscript, 2008) indicates that Georges Creek experiences larger unit peaks and shorter centroid lags than the adjacent, forested Savage River watershed. As the use of an adjacent control watershed normalizes for climatic variability, differences in stormflow are attributed to the effect of surface mining and reclamation. However, these analyses have not investigated the effect of mining on flood response as an explicit function of the proportion of a watershed covered by mining and subsequent reclamation.

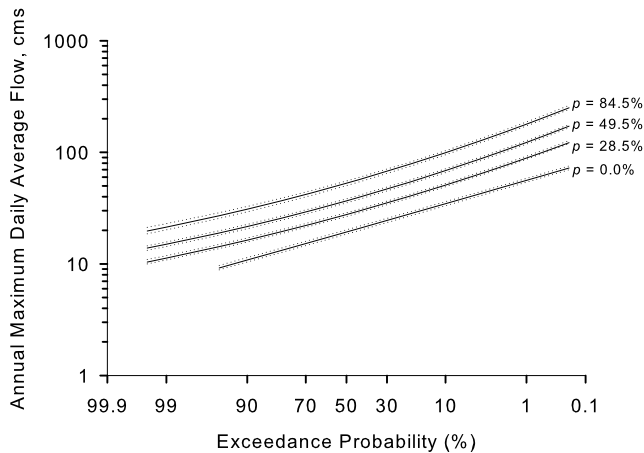
[35] To explore the impact of surface mine activities beyond what already exists in Georges Creek, we used the Hydrological Simulation Program-Fortran (HSPF) as a modeling framework. Surface mining and reclamation present at least two major challenges to the calibration of basin-scale hydrologic models. First, reclaimed sites are difficult to identify using remote sensing imagery, and our modeling would not have been possible without the enhanced LULC classification provided by Townsend *et al.* [2009]. In addition, the deep shaft mining tunnels in Georges Creek are common for watersheds across the CAP. We modeled the effect of underground mine tunnels and deep shaft drainage in Georges Creek as a delayed storage mechanism, which showed superior calibration results relative to the standard treatment of deep aquifer losses using the HSPF parameter DEEPPFR. This approach may be of use for future analyses as many surface-mined watersheds in the CAP have legacy effects due to prior underground coal mining [Atkins *et al.*, 2005].

[36] We also introduced to our calibration elements of the GLUE technique [Beven and Binley, 1992] as well as the Gauss-Marquardt-Levenburg (GML) method for numerical search of objective function minima [Levenburg, 1944; Marquardt, 1963]. All calculations were performed using freely available PEST and PD\_MS2 software [Gallagher

**Table 6.** Mean Peak (Daily) Stormflow Errors for Simulations Using Parameters Calibrated for Each Simulation Time Period<sup>a</sup>

Mean Peak Storm Errors (%), 10-Year Simulation Period	Using Parameters From Calibration Time Period		
	1971–1981	1982–1992	1994–2004
1971–1981	<b>-30.5</b> (-31.9 to -29.2)	<b>-26.3</b> (-29.2 to -23.3)	<b>-29.0</b> (-31.3 to -26.7)
1982–1992	<b>-15.0</b> (-16.4 to -13.7)	<b>-11.5</b> (-13.0 to -10.0)	<b>-12.4</b> (-15.2 to -9.5)
1994–2004	<b>-15.3</b> (-17.0 to -13.6)	<b>-15.8</b> (-17.6 to -14.0)	<b>-14.5</b> (-16.9 to -12.0)

<sup>a</sup>Bold values are for statistics using  $\bar{q}_t$  time series (equation (7)); values in parentheses indicate predictive uncertainty intervals.



**Figure 6.** Representative log Pearson III flood frequency curves for Georges Creek under four proportions of mining/reclamation:  $p = 0.0\%$ ,  $28.5\%$ ,  $49.5\%$ , and  $84.5\%$ . Solid lines represent plots created using mean, simulated flows (equation (7)), and dotted lines represent plots created using upper and lower predictive uncertainty intervals. Largest probabilities for  $p = 0.0\%$  were not plotted, as the values fell below the cutoff threshold of the program PEAKFQ.

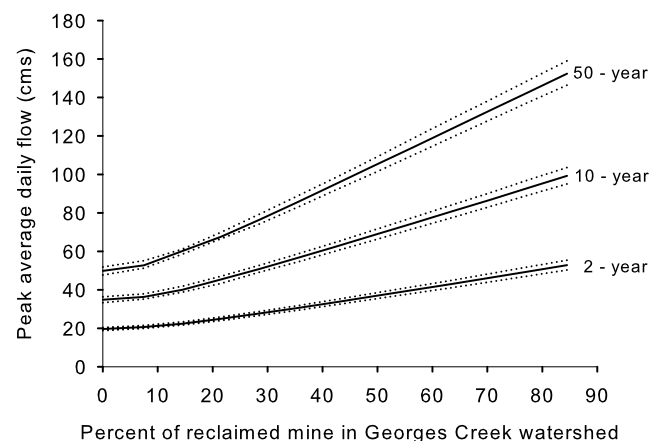
and Doherty, 2007]. Through coupled use of both the GLUE and GML techniques, we were able to obtain robust parameterizations with relatively narrow predictive uncertainty intervals. This approach, and variations thereof [e.g., Doherty and Johnston, 2003], show advantages over pure use of the GLUE technique, as the guided GML optima search yielded multiple parameter sets for each calibration that had only minor differences in model performance (and hence narrow predictive uncertainty bounds), while consistently delivering objective function values below those obtained during Monte Carlo simulations alone.

[37] Comparing calibrations for Georges Creek across three different time periods, it was anticipated that calibration for a given time period would yield the best values for Nash-Sutcliffe efficiency and stormflow volume and peak flow errors (on-diagonal elements, Tables 4, 5, and 6, respectively). This was not necessarily the case, however, as the statistics compared were not the equivalent of the objective function used for calibration. The mean peak flow errors for the 1971–1981 time period (Table 6, top row) were consistently double that of other time periods. This may have been due in part to the fact that only 30% of the days for which precipitation occurred during the 1971–1981 simulations had hourly data to guide disaggregation (as compared to 56% and 62% for the 1982–1992 and 1994–2004 time periods). Remaining daily precipitation records were disaggregated using a triangular distribution, which could dampen model response by reducing rainfall intensities and hence peak flows (on an hourly basis). More important to our modeling, which reports statistics on daily flows, triangular distributions can incorrectly distribute precipitation across 2 calendar days, thus minimizing simulated peaks relative to observed (Table 6) while maintaining overall storm mass balances (Table 5).

[38] Two important findings can be drawn from our LULC change analysis. First, for a given return interval, flood magnitude increases linearly ( $r^2 > 0.99$ ) with increasing

percentages of watershed area  $p$  affected by surface mining and subsequent reclamation. Second, larger rates of increase are expected for lower return intervals. These results highlight the unique nature of surface-mined lands relative to other LULC. For example, recent empirical work based on long-term data in the CAP indicates timber harvest has relatively negligible effects on flood magnitude [Kochenderfer et al., 2007]. While deforestation affects interception storage and evapotranspiration, producing increases in water yield [Bosch and Hewlett, 1982], forest soils are generally not severely compacted during logging operations and retain a considerable capacity to absorb precipitation and mitigate floods. Thus disturbances that involve vegetation removal alone (e.g., timber harvest) may alter water yields but are unlikely to affect flood frequency or magnitude in the CAP. Reclaimed surface mines, which are also cleared of natural vegetation, are subjected to massive soil compaction as a consequence of using large earthmovers to complete grading operations in compliance with the Surface Mine Control and Reclamation Act of 1977 [U.S. Congress, 1977]. Although the reclaimed mines may meet contour guidelines, the disturbed/reconstructed soils are often a poor medium for plant growth [Simmons et al., 2008] and unfavorable for reforestation [Bussler et al., 1984].

[39] Though not identical, our results more closely follow the trends reported for a different form of LULC: urbanization. In a synthesis of published empirical data on the effect of urbanization on flood response, Hollis [1975] showed that minor levels of urbanization (e.g., 20% impervious surfaces in a watershed) had a much larger effect on short return interval floods than on the magnitude of large, infrequent floods. While we did not find this same pattern for reclaimed lands, Hollis [1975] also showed that floods with long return intervals (e.g., 100-year events) may be doubled in size when impervious cover shifted from 20% to 30% of a basin. Our results do support a similar expectation as mine reclamation reaches higher proportions of watershed area. In another study of changing flood response with varying amounts of urbanization, Wissmar et al. [2004] showed that the rate of increase of 50-year flood events was double that of 10-year



**Figure 7.** Peak mean daily flows for 2-, 10-, and 50-year flood events versus the proportion of Georges Creek affected by reclaimed mining. Plot assumes that the percentages of active mine, agriculture, urban, and water are constant at 1999 values (Table 1).

events (as a function of the amount of impervious cover in a watershed), which generally supports our finding of greater rates of increase for longer return interval events.

[40] The dominant mechanism affecting stormflows from urban areas is typically attributed to impervious surfaces (roads, parking lots, buildings, etc.). Urban areas more often than not use civil engineering structures such as storm sewers and retention ponds to mitigate peak flow events that are exacerbated by large areas of impervious cover. Reclaimed surface mines are not impervious, but the accumulation of large areas of severely compacted soils with infiltration rates an order of magnitude less than nearby, undisturbed sites yields a surface that is “almost impervious.” Unlike urbanization, surface mine reclamation rarely uses permanent civil engineering structures, as this is not required by law for mitigation of flood response. Rather, a return to appropriate contours and revegetation are considered sufficient for compliance with federal mine reclamation standards. While these strategies may be useful for returning aesthetics and other important qualities to previously mined sites, they do not appear to be effective at restoring key features of the hydrologic regime.

## 5. Conclusions

[41] Surface mining of bituminous coal is a major vector of change for the central Appalachian Plateau. Our findings suggest that the effect of surface mining (and reclamation) at the scale of Georges Creek watershed (187 km<sup>2</sup>) has interesting parallels to what would be expected for urban areas with sizable proportions of impervious area. Although reclaimed surface mines have nonzero infiltration rates, the rates are typically an order of magnitude (or 2) smaller than for undisturbed forest. In addition, the flood response from impervious urbanized areas is often mitigated by engineered storm water management structures that are typically absent on reclaimed mines. As a consequence, the act of mine reclamation should not be interpreted as meaning the land is returned to a state that is the hydrological equivalent of the premining landscape. This has been previously demonstrated empirically at the small watershed scale [e.g., Bonta et al., 1997; Negley and Eshleman, 2006]. Our work demonstrates the ramifications of surface mining and reclamation at the scale of larger catchments (100–1000 km<sup>2</sup>). Although our results are specific to Georges Creek, the trends observed are expected to be similar across watersheds in the CAP, thus offering some guidance as to possible shifts in flood regimes as a consequence of future surface mining activities.

[42] **Acknowledgments.** We gratefully acknowledge financial support for this work from the NASA Land Cover Land Use Change (LCLUC) program, grant NNG06GC83G to the University of Maryland Center for Environmental Science, Appalachian Laboratory. This paper is scientific contribution 4249 from the University of Maryland Center for Environmental Science. We are particularly grateful to the three anonymous reviewers for critical reviews that greatly improved the manuscript.

## References

- Allan, J. D., D. L. Erickson, and J. Fay (1997), The influence of catchment land use on stream integrity across multiple spatial scales, *Freshwater Biol.*, *37*, 149–161, doi:10.1046/j.1365-2427.1997.d01-546.x.
- Atkins, J. T., J. B. Wiley, and K. S. Paybins (2005), Calibration parameters used to simulate streamflow from application of the Hydrologic Simulation Program-FORTRAN model (HSPF) to mountainous basins containing coal mines in West Virginia, *U.S. Geol. Surv. Sci. Invest. Rep.*, 2005-5099, 82 pp.
- Beven, K. J. (2006), A manifesto for the equifinality thesis, *J. Hydrol.*, *320*, 18–36.
- Beven, K. J., and A. Binley (1992), The future of distributed models; model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298, doi:10.1002/hyp.3360060305.
- Bicknell, B. R., J. C. Imhoff, J. L. Kittle, T. H. Jobs, and A. S. Donigian (2001), Hydrological Simulation Program-FORTRAN Version 12 user's manual, Aqua Terra Consult., Mountain View, Calif. (Available at <http://www.epa.gov/waterscience/basins/bnsdocs.html>)
- Bonta, J. V., C. R. Amerman, T. J. Harlukowicz, and W. A. Dick (1997), Impact of coal surface mining on three Ohio watersheds—Surface water hydrology, *J. Am. Water Resour. Assoc.*, *33*, 907–917, doi:10.1111/j.1752-1688.1997.tb04114.x.
- Bosch, J. M., and J. D. Hewlett (1982), A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration, *J. Hydrol.*, *55*, 3–23, doi:10.1016/0022-1694(82)90117-2.
- Brun, S. E., and L. E. Band (2000), Simulating runoff behavior in an urbanizing watershed, *Comput. Environ. Urban Syst.*, *24*, 5–22, doi:10.1016/S0198-9715(99)00040-X.
- Burges, S. J., M. S. Wigmosta, and J. M. Meena (1998), Hydrological effects of land-use change in a zero-order catchment, *J. Hydrol. Eng.*, *3*, 86–97, doi:10.1061/(ASCE)1084-0699(1998)3:2(86).
- Bussler, B. H., W. R. Byrnes, P. E. Pope, and W. R. Chaney (1984), Properties of minesoil reclaimed for forest land use, *Soil Sci. Soc. Am. J.*, *48*, 178–184.
- Chong, S. K., M. A. Becker, S. M. Moore, and G. T. Weaver (1986), Characterization of mined land with and without topsoil, *J. Environ. Qual.*, *15*, 157–160.
- Coon, W. F. (2003), Simulating land-use changes and stormwater-detention basins and evaluating their effect on peak streamflows and stream-water quality in Irondequoit Creek Basin, New York; a users manual for HSPF and GENSCN, *U.S. Geol. Surv. Open File Rep.*, 03–136, 29 pp.
- Criss, R. E., and W. E. Winston (2008), Do Nash values have value? Discussion and alternate proposals, *Hydrol. Processes*, *22*, 2723–2725, doi:10.1002/hyp.7072.
- Doherty, J. (2005), PEST: Software for model-independent parameter estimation, Watermark Numer. Comput., Brisbane, Australia. (Available at <http://www.sspa.com/pest>)
- Doherty, J. (2007), PEST surface water utilities, Watermark Numer. Comput., Brisbane, Australia. (Available at <http://www.sspa.com/pest/download/swutils.pdf>)
- Doherty, J., and J. M. Johnston (2003), Methodologies for calibration and predictive analysis of a watershed model, *J. Am. Water Resour. Assoc.*, *39*, 251–265, doi:10.1111/j.1752-1688.2003.tb04381.x.
- Eshleman, K. N. (2004), Hydrological consequences of land use change: A review of the state-of-science, in *Ecosystems and Land Use Change, Geophys. Monogr. Ser.*, vol. 153, edited by R. Defries, G. Asner, and S. Houghton, pp. 13–29, AGU, Washington, D. C.
- Ewen, J., G. O'Donnell, A. Burton, and E. O'Connell (2006), Errors and uncertainty in physically based rainfall-runoff modelling of catchment change effects, *J. Hydrol.*, *330*, 641–650, doi:10.1016/j.jhydrol.2006.04.024.
- Flynn, K. M., W. H. Kirby, and P. R. Hummel (2006), User's manual for program PeakFQ, annual flood-frequency analysis using bulletin 17B guidelines, *U.S. Geol. Surv. Tech. Methods, Book 4, Chap. B4*, 42 pp.
- Gallagher, M., and J. Doherty (2007), Parameter estimation and uncertainty analysis for a watershed model, *Environ. Model. Softw.*, *22*, 1000–1020, doi:10.1016/j.envsoft.2006.06.007.
- Gutierrez-Magness, A. L. (2005), A strategy for calibrating the HSPF model, Ph.D. dissertation, Dep. of Civ. and Environ. Eng., Univ. of Md., College Park.
- Hollis, G. E. (1975), The effect of urbanization on floods of different recurrence interval, *Water Resour. Res.*, *11*, 431–435, doi:10.1029/WR011i003p00431.
- Homer, C., C. Q. Huang, L. M. Yang, B. Wylie, and M. Coan (2004), Development of a 2001 national land-cover database for the United States, *Photogramm. Eng. Remote Sens.*, *70*, 829–840.
- Hornbeck, J. W., R. S. Pierce, and C. A. Federer (1970), Streamflow changes after forest clearing in New England, *Water Resour. Res.*, *6*, 1124–1132, doi:10.1029/WR006i004p01124.
- Hummel, P., J. Kittle, and M. Gray (2001), WDMUTIL 2.0: A tool for managing watershed modeling time-series data, users manual, Aqua Terra Consult., Decatur, Ga. (Available at <http://www.epa.gov/waterscience/basins/bnsdocs.html>)

- Interagency Advisory Committee on Water Data (1982), Guidelines for determining flood flow frequency, *Bull. 17B*, Hydrol. Subcomm., Off. of Water Data Coord., U.S. Geol. Surv., Reston, Va.
- Jacomino, V. M. F., and D. E. Fields (1997), A critical approach to the calibration of a watershed model, *J. Am. Water Resour. Assoc.*, 33, 143–154, doi:10.1111/j.1752-1688.1997.tb04091.x.
- Jia, Y., and T. B. Culver (2008), Uncertainty analysis for watershed modeling using generalized likelihood uncertainty estimation with multiple calibration measures, *J. Water Resour. Plann. Manage.*, 134, 97–106, doi:10.1061/(ASCE)0733-9496(2008)134:2(97).
- Jorgensen, D. W., and T. W. Gardner (1987), Infiltration capacity of disturbed soils: Temporal change and lithologic control, *Water Resour. Bull.*, 23, 1161–1172.
- Kochenderfer, J. N., M. B. Adams, G. W. Miller, and J. D. Helvey (2007), Factors affecting large peakflows on Appalachian watersheds: Lessons from the Fernow experimental forest, *Res. Pap. NRS-3*, 24 pp., U.S. Dep. of Agric. For. Serv. North. Res. Stn., Newtown Square, Pa. (Available at <http://www.nrs.fs.fed.us/pubs/2838>)
- Krause, P., D. P. Boyle, and F. Base (2005), Comparison of different efficiency criteria for hydrological model assessment, *Adv. Geosci.*, 5, 89–97.
- Levenburg, K. (1944), A method for solution of certain problems in least squares, *Q. Appl. Math.*, 2, 164–168.
- Loveland, T., et al. (2003), Land-use/land cover change, in *Strategic Plan for the Climate Change Science Program Final Report*, edited by J. R. Mahoney, pp. 63–70, U.S. Clim. Change Program, Washington, D. C.
- Lumb, A. M., R. B. McCammon, and J. L. Kittle (1994), Users manual for an expert system (HSPexp) for calibration of the Hydrologic Simulation Program-FORTRAN, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, 94–4168, 102 pp. (Available at <http://water.usgs.gov/software/HSPexp/code/doc/hspexp.pdf>)
- Marquardt, D. (1963), An algorithm for least-squares estimation of non-linear parameters, *SIAM J. Appl. Math.*, 11, 431–441, doi:10.1137/0111030.
- Maryland Department of Natural Resources (2001), Georges Creek watershed characterization, 87 pp., Annapolis. (Available at [http://dnrweb.dnr.state.md.us/download/bays/ger\\_char.pdf](http://dnrweb.dnr.state.md.us/download/bays/ger_char.pdf))
- Negley, T. L., and K. N. Eshleman (2006), Comparison of stormflow responses of surface-mined and forested watersheds in the Appalachian Mountains, USA, *Hydrol. Processes*, 20, 3467–3483, doi:10.1002/hyp.6148.
- Office of Surface Mining (2004), Office of Surface Mining annual report: 2004, U.S. Dep. of the Interior, Washington, D. C.
- Pederson, T. A., A. S. Rogowski, and R. Pennock Jr. (1980), Physical characteristics of some minesoils, *Soil Sci. Soc. Am. J.*, 44, 321–328.
- Potter, K. W. (1991), Hydrological impacts of changing land management practices in a moderate sized agricultural catchment, *Water Resour. Res.*, 27, 845–855, doi:10.1029/91WR00076.
- Potter, K. N., F. S. Carter, and E. C. Doll (1988), Physical properties of constructed and undisturbed soils, *Soil Sci. Soc. Am. J.*, 52, 1435–1438.
- Samuel, J. M., and M. Sivapalan (2008), Effects of multiscale rainfall variability on flood frequency: Comparative multisite analysis of dominant runoff processes, *Water Resour. Res.*, 44, W09423, doi:10.1029/2008WR006928.
- Simmons, J. A., W. S. Currie, K. N. Eshleman, K. Kuers, S. Monteleone, T. L. Negley, B. R. Pohlrad, and C. L. Thomas (2008), Forest to reclaimed mine land use change leads to altered ecosystem structure and function, *Ecol. Appl.*, 18, 104–118, doi:10.1890/07-1117.1.
- Townsend, P. A., D. P. Helmers, C. C. Kingdon, B. E. McNeil, K. M. de Beurs, and K. N. Eshleman (2009), Changes in the extent of surface mining and reclamation in the Central Appalachians: 1976–2006, *Remote Sens. Environ.*, 113, 62–72, doi:10.1016/j.rse.2008.08.012.
- U.S. Congress (1977), Surface Mining Control and Reclamation Act of 1977, *Public Law 95–87*, 95th Congress, 1st session.
- U.S. Environmental Protection Agency (2000), Estimating hydrology and hydraulic parameters for HSPF, BASINS Technical Note 6, *EPA-823-R00-012*, Off. of Water, Washington, D. C. (Available at <http://www.epa.gov/waterscience/basins/bnsdocs.html>)
- U.S. Environmental Protection Agency (2001), Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) version 3.0 Users Manual, *EPA-823-H-01-001*, Washington, D. C.
- Vogelmann, J. E., S. M. Howard, L. Yang, C. E. Larson, B. K. Wylie, and J. V. Van Driel (2001), Completion of the 1990's National Land Cover Data Set for the conterminous United States, *Photogramm. Eng. Remote Sens.*, 67, 650–662.
- Wissmar, R. C., R. K. Kim, and M. G. Logsdon (2004), Effects of changing forest and impervious land covers on discharge characteristics of watersheds, *Environ. Manage. N. Y.*, 34, 91–98, doi:10.1007/s00267-004-0224-5.

---

K. N. Eshleman, J. R. Ferrari, T. R. Lookingbill and B. McCormick, Appalachian Laboratory, University of Maryland Center for Environmental Science, 301 Braddock Road, Frostburg, MD 21532, USA. (jfferrari@al.umces.edu)

P. A. Townsend, Russell Laboratories, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI 53706, USA.