Effects of Multi-gage Calibration on the Location of Critical Source Areas Using SWAT

# Abstract

Ecosystem services (ES) are the benefits humans receive from nature. In recent years payment schemes have been devised to incentivize land owners to manage their lands in ways that will promote the conservation and restoration of ecosystem services. In order to ensure land and water managers devote funds effectively they need to know the locations of critical areas which should be targeted. Models are used to locate these Critical Source Areas (CSAs), however there are many models available, and each have different algorithms. Recent research has shown that using an uncalibrated watershed model reveals the same CSAs as a lumped calibrated model, but that different models calibrated at a single gage reveal slightly different CSAs. Since the location of CSAs is inherently spatial, we hypothesize that multi-site calibration will significantly alter the final lumped model calibration, and will therefore also alter the location of CSA’s throughout the watershed. We use the Soil and Water Assessment Tool (SWAT) to develop a model for two sub-basins in Northwest Oregon, USA. The first sub-basin, Tualatin, is mixed use and consists of forest, agriculture, and urban land uses. It is highly managed and therefore data rich. The second basin, Yamhill, is dominated by forest and agricultural lands. It has some detailed flow monitoring, but water quality monitoring is sparse in comparison. We compare the uncalibrated CSA’s for each basin with their calibrated locations. We then compare how the calibrated results for the Tualatin with one gages compares to the results of calibrating the Tualatin with multiple gages.

# 1. Introduction

Many of earth’s biophysical processes provide the foundation upon which the global economy is based (MA 2005). Yet the services we receive from nature, termed ecosystem services, are economically undervalued (Costanza et al 1997). As a result, many economic decisions do not consider the impacts human activities will have on the ecosystems we depend upon. In the absence of strict government regulations promoting healthy land and water management, researchers are exploring innovative solutions such as providing payments to landowners who manage their lands using sustainable methods. This approach is termed the Payment for Ecosystem Services (PES) approach (Farley and Costanza 2010). In northwest Oregon, the Oregon Department of Environmental Quality (ODEQ) developed a payment scheme to enable Clean Water Services (CWS), the owner of wastewater treatment facilities in the Tualatin basin, to offset thermal loads by planting trees instead of constructing a refrigeration unit to cool water (Cochran and Logue 2011). The Willamette Partnership (<http://willamettepartnership.org/>) was created in 2004 and has developed a water quality trading program to help improve water quality throughout the Willamette River watershed.

In order to ensure these payments are directed to the areas of the watershed that can provide the greatest improvement to water quality, researchers use both field studies and models. While field studies are the most accurate, they are also the most resource intensive. After a thorough search for empirical studies of erosion off the landscape in Washington county, all we could find was a Washington County Soil & Water conservation district project using the EPIC model to predict erosion off agricultural lands (Moberg 1995). While the report discusses the need for further field scale studies to validate the model and mentions funding for future work, the results of that project were inconclusive and ultimately went unpublished. Therefore, while there is a lot of uncertainty in modeled results, it is often times the only information available.

Given this reality, it is important to understand how the choice and calibration of models determines the identification of CSAs. Niraula et al (2011) compare an uncalibrated SWAT model to a lumped calibrated SWAT model and found no significant difference between the identification of CSA’s in model output. Niraula et al (2013) compare a calibrated SWAT model with a calibrated GWLF model and found that due to differences in model algorithms, CSAs can vary depending on the model used. These findings are important and suggest that further research must be done to better understand how model choice and calibration affects the location of CSAs. In both of these studies, one gage was used to calibrate the model. While one gage may be sufficient for studies interested in loadings at the mouth of the basin, and while often times only one gage is available for calibration, the spatial nature of CSA identification means that important information may be lost by calibrating with only one gage. Chiang et al (2012) show that calibration results for total nitrogen improve with the inclusion of multiple gages, but they do not look at how these calibrations change the locations of CSAs. In this study we investigate the following research questions: 1) Do the locations of CSA’s change when models of the Yamhill and Tualatin basins are calibrated with one gage for each basin. 2) Do the locations of CSA’s in the Tualatin change when the model is calibrated using local calibration with three gages, and 3) If the locations do change, why do they change and what are the spatial patterns of these changes?

# 2. Study Site

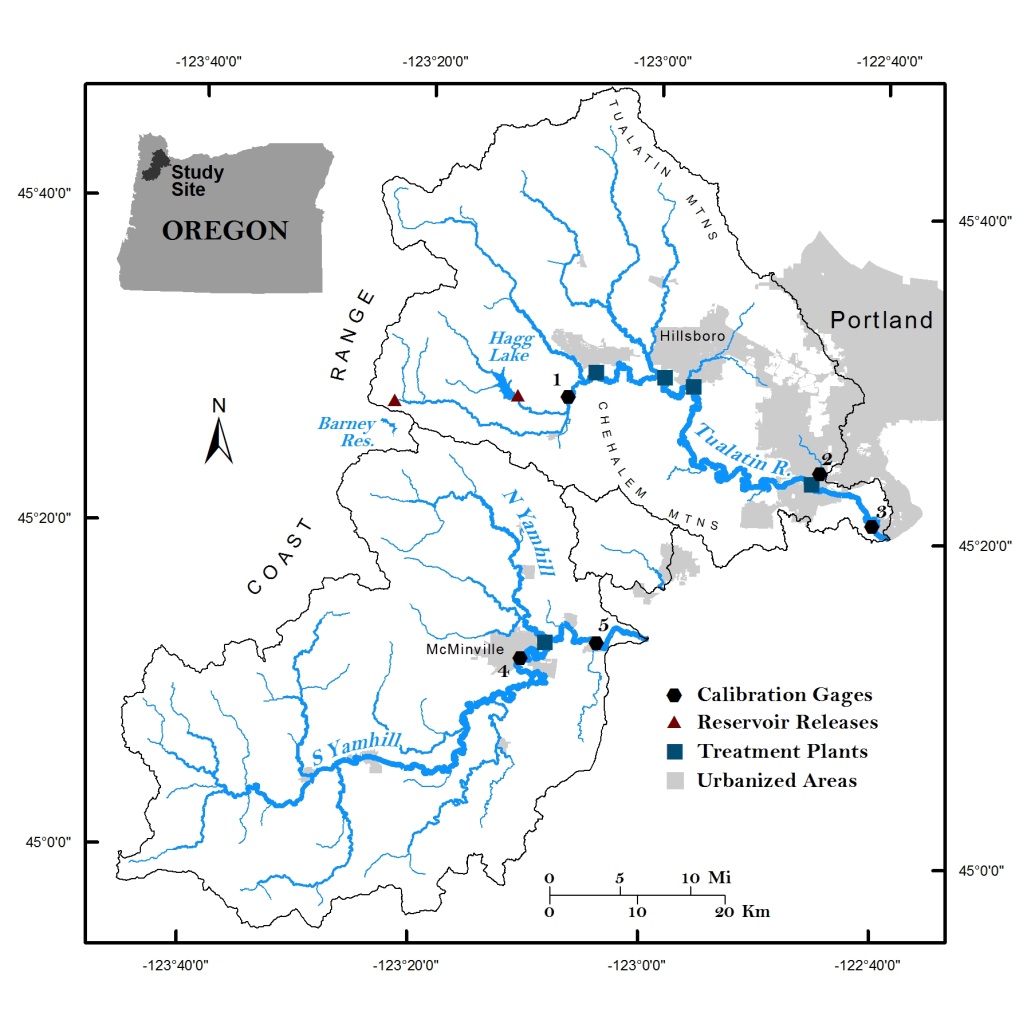
## 2.1 Tualatin

The 1,829 km2 Tualatin River Basin roughly shares the borders of Washington County in Northwestern Oregon (Fig. 1). The basin is bordered by the Coast range to the west, Tualatin Mountains (West Hills) to the north and east, and the Chehalem Mountains to the south. With the exception of its headwaters which originate in the Coast Range, the Tualatin River is a low-gradient, meandering river which travels 130 km east, before emptying into the Willamette River. Elevation in the basin ranges from a high of 1,057 m to a low of 17 m at the river’s mouth, and has a mean elevation of 195 m. Soils in the basin formed from weathering of the Columbia River Basalt Group, and deposition of the Willamette Silts by the Missoula Floods during the late Pleistocene. The region has a modified marine climate, dominated by cool wet winters, and warm dry summers (Fig 2). In upper elevations, annual precipitation ranges from 1,330 to 3,280 mm and average daily temperatures range from 4 to 27°C in the summer and -16 to 12°C in the winter. In the valley, annual precipitation ranges from 740 to 1,850 mm, and average daily temperatures range from 10 to 31°C in the summer, and -10 to 15°C in the winter.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 1**: Historic water budget (mm) for the time period 09/30/1994 - 10/22/2010 | | | | | |
| **Basin** | **Precip** | **Runoff** | **PET** | **AET** | **Runoff Ratio** |
| ***Tualatin*** | 1,266 | 808 | 642 | 458 | 0.64 |
| ***Yamhill*** | 1,585 | 865 | 655 | 720 | 0.55 |

Stream flow is largely precipitation dominated with peak flows occurring during the winter wet season, October through May, and low flows occurring during the summer dry months, June through August (Fig 2). The basin has a runoff ratio of 0.64 (Table 1). Two large dams alter the hydrology of the basin. Scoggins Dam on Scoggins Creek provides supplemental flows of around 211 cfs in the summertime as well as recreational opportunities for local residents. Barney reservoir provides additional flows of around 14 cfs as an inter-basin water transfer from the Trask River to the upstream portion of the Tualatin. Clean Water Services (CWS) operates four waste-water treatment plants (WWTPs) located along the main stem of the Tualatin River. The two downstream plants, Durham and Rock Creek, process the majority of effluent, while the two upstream plants, Hillsboro and Forest Grove, maintain reserve capacity for anticipated population growth.

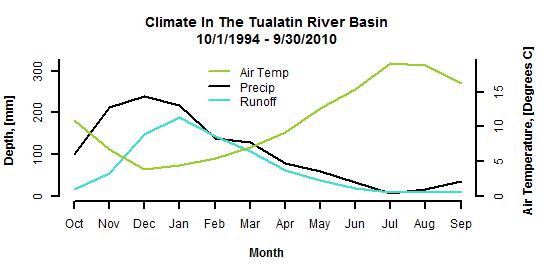
Agricultural land dominates the basin. Approximately 49% of land in the basin is cultivated, while forested lands comprise 23%, and 14% has been developed. The majority of the basin (93%) is privately owned. Of public lands, 5% is owned by the State of Oregon and 2% is owned by the Bureau of Land Management (ODEQ 2001).

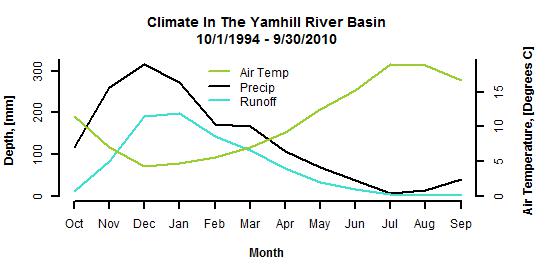


**Figure 1:** Map of the Tualatin and Yamhill River basins. Gage numbers are referenced in Table 2.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 2:** Gages used for calibration. F = Flow, TSS = Total Suspended Solids, TN = Total Nitrogen, TP = Total Phosphorus. | | | | |
| ***Gage #*** | ***Name*** | ***Organization*** | ***ID #*** | ***Constituents*** |
| **1** | Tualatin River at Dilley | USGS/CWS | 14203500 | F, TSS, TN, TP |
| **2** | Fanno Creek at Durham | USGS/CWS | 14206950 | F, TSS, TN, TP |
| **3** | Tualatin River at West Linn | USGS/CWS | 14207500 | F, TSS, TN, TP |
| **4** | South Yamhill River at McMinnville | USGS | 14194150 | F |
| **5** | Yamhill Water Quality Station | DEQ | 10363 | S, TN, TP |

Due to agriculture, timber harvesting, and rapid urbanization in the mid-20th century, the basin suffered from poor water quality. In 2001, EPA approved the first TMDLs for Temperature, Bacteria, Dissolved Oxygen, pH, and Phosphorus (ODEQ 2001). Changes have been made to the TMDLs over the years as needs have arisen, and water quality has improved. However, some rapidly urbanizing areas of the basin still experience water quality problems (Boeder and Chang 2008). CWS is the designated management agency in the basin, and is in charge of monitoring and implementing water quality management plans. Climate change studies in the region indicate that rising air temperatures will accentuate the seasonal range of stream flows, with flows expected to increase in the winter and decrease in the summer (Hamlet and Lettenmaier 1999; Franczyk and Chang 2009; Chang and Jones 2010; Praskievicz and Chang 2011).





**Figure 2:** Climate in the Tualatin and Yamhill sub-basins

2.2 Yamhill

The Yamhill sub-basin lies to the south of the Tualatin, and drains 1,998 km2 (Figure 1). The two main rivers, North and South Yamhill, flow southeast and northeast, respectively, until they converge and flow east before emptying into the Willamette River. Elevation in the basin ranges from 1,084 m in the Coast Range to 18 m at the mouth of the Yamhill and has a mean elevation of 217 m. Soils in the basin have similar provenance to those in the Tualatin. Annual precipitation ranges from 1,560 to 3,880 mm in high elevations and 560 to 1,710 mm in lower elevations. Average daily temperatures at high elevations range from -14 to 12 degrees in the winter and 7 to 27 degrees in the summer. Low elevation daily temperatures range from -10 to 15 degrees in the winter and 10 to 30 degrees in the summer.

The Yamhill River system is significantly less managed than the Tualatin. There is no major reservoir in the Yamhill to supplement flows or provide flood control, so during the summer measured flows have dropped to as little as 1.4 cfs, while winter wet seasons have seen flows as large as 40,300 cfs. The runoff ratio is 0.55 (Table 1). There is also less development in the basin. 40% of the basin is forested. One third of the basin consists of cultivated crops, and only 7% is developed.

# 3 Data and Methods

## 3.1 Data

A 30 m resolution digital elevation model (DEM) and a vectorized stream network were downloaded for our study area from the NHDPlus (version 1) website (NHD Plus 2006). We downloaded the 30 m resolution, 2006 National Land Cover Dataset (NLCD) produced by MRLC (2011), and used the state level soils dataset STATGO for our analysis. While STATGO has a coarser resolution than the SSURGO dataset, SSURGO has a data gap in a portion of the study site. Furthermore, Mukunden et al (2010) found that there is no statistical difference between the two for model accuracy of flow and sediment. We downloaded daily interpolated climate grids for the years 1979 to 2010 from the Interactive Numeric & Spatial Information Data Engine (Abatzoglou 2013), which we used for model calibration. Water quantity and quality data was acquired from a combination of USGS gaging stations and CWS gaging stations throughout the Tualatin. WWTP discharges as well as Barney Reservoir and Hagg Lake releases were also acquired from CWS. Site level specifications of Scoggins dam were compiled from Bureau of Reclamation and USGS reports.

The availability of water quantity and quality data collection varies substantially. Yamhill has relatively limited data compared to the Tualatin. In both basins, the temporal resolution of water quality measurements is typically around two to five measurements per month. In order to develop a more complete time series of sediment, nitrogen and phosphorus loads, we use the LOADEST software (Runkel et al 2004) (Table 2) to estimate a continuous daily time series. These loads were then aggregated to monthly loads for model calibration.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3**:** LOADEST sediment and nutrient results. | | | | | | | |
| **Parameter** | **Calibration Period** | **Grab Samples** | | **Estimation Period** | **NSE** | **PBIAS** | **R2** |
| *Tualatin River at West Linn* | | | | | | | |
| ***TSS*** | 1988 – 2010 | 828 | | 1981 – 2010 | 0.06 | 11.14 | 0.94 |
| ***TN*** | 1974 – 2002 | 545 | | 1981 – 2010 | 0.93 | 0.12 | 0.95 |
| ***TP*** | 1974 – 2010 | 972 | | 1981 – 2010 | 0.65 | 5.63 | 0.92 |
| *Tualatin River at Dilley* | | | | | | | |
| ***TSS*** | 1984 – 2011 | 1014 | | 1981 – 2010 | 0.45 | -10.24 | 0.89 |
| ***TN*** | 1984 – 2011 | 1,007 | | 1981 – 2010 | 0.68 | 3.15 | 0.88 |
| ***TP*** | 1984 – 2011 | 1,032\* | | 1981 – 2010 | 0.67 | -2.80 | 0.84 |
| *Fanno Creek at Durham* | | | | | | | |
| ***TSS*** | 10/1/1993 – 2012 | 530 | | 10/1/1993 – 2010 | -1.76 | 46.09† | 0.93 |
| ***TN*** | 10/1/1993 – 2012 | 623\*\* | | 10/1/1993 – 2010 | 0.93 | 2.72 | 0.98 |
| ***TP*** | 10/1/1993 – 2012 | 733 | | 10/1/1993 – 2010 | 0.66 | 1.75 | 0.98 |
| *Yamhill Water Quality Station* | | | | | | | |
| ***TSS*** | 10/1/1994 – 2012 | 164 | | 10/1/1994 – 2010 | 0.72 | 3.58 | 0.96 |
| ***TN*** | 10/1/1994 – 2007 | 132 | | 10/1/1994 – 2010 | 0.77 | 8.86 | 0.97 |
| ***TP*** | 10/1/1994 – 2012 | 164 | | 10/1/1994 – 2010 | 0.85 | -1.74 | 0.97 |
| \* 25 samples registered below hardware detection limits  \*\* 1 sample registered below hardware detection limits | | | † LOADEST does not recommend using models with PBIAS greater than 25% | | | | |

## 3.2 Methods

### 3.2.1 SWAT Model

SWAT is a physically based, daily time-step model. It accounts for both terrestrial and in-stream processes (Fig 3). To model flow, SWAT uses either the SCS curve number procedure (SCS, 1972), or the Green & Ampt infiltration method (1911). While the Green & Ampt method is more robust, it requires sub-daily precipitation. Since hourly precipitation is not available for our study region, we use the SCS curve number approach.

**Climate**

**Topography**

**Soil Structure**

**Land Use**

**Hydrologic Response Unit (HRUs)**

**Sub-basins**

**Terrestrial Component:**

***Water Yield*** (SCS Curve #)

***Sediment Yield*** (MUSLE)

***Nutrient Yield*** (Mass

Balance)

**In-stream** **Processes**

**Sub-basin/reach scale**

Adds terrestrial yields, sub-surface contributions, and inputs from upstream reaches to current reach, and computes in-stream processes for current reach.

**Watershed**

**Outputs**

***Flow***

***Sediment***

***Total Nitrogen***

***Total Phosphorus***

**Figure 3**: Conceptual diagram of SWAT.

SWAT uses the Modified Universal Soil Loss Equation (MUSLE, Williams, 1975) to model sediment transport across the landscape. The MUSLE uses runoff instead of precipitation as a measure of erosive energy. Since the majority of field level studies have been done using the USLE, SWAT also provides USLE output for comparison purposes. The nitrogen mass balance is budgeted into five pools, and two main categories: Mineral N and Organic N. Mineral N consists of the ammonia and nitrate pools, while organic N consists of the fresh organic N (biomass) and active and stable organic N pools. The Phosphorus mass balance is budgeted into six pools split between mineral and organic P. Mineral P consists of the stable, active, and solution pools, while organic P consists of the stable, active, and fresh (biomass) pools (Neitsch 2011).

There are several options for modeling in-stream sediment processes. Channel sediment deposition and re-entrainment are modeled using one of four equations. The Simplified Bangold equation is the default, but does not pool sediment by particle size. Three other options are available to do so. We chose the Simplified Bangold equation since the two basins have relatively homogenous sediments.

In-stream nutrient processes are optional, but are sometimes necessary for accurate results. SWAT uses algorithms from the QUAL2E model to simulate nutrient transport between the stream and algae (Brown and Barnwell, 1987).

### 3.2.2 Calibration Metrics

The efficacy of a watershed model is measured by how closely the model’s simulated output compares to measured observed flow data. Among the many metrics, we used the three suggested by Moriasi (2007): Nashe-Sutcliffe Efficiency (NSE), percent bias (PBIAS), and the RMSE-observations standard deviation (RSR).

The NSE is calculated as follows:

where represents the number of observations, is the observed data point, is the simulated data point, and is the mean of all the observed data points. If the model perfectly fits the observed data, . If the model is just as good as taking the mean of the observed data, . If , you are better off using the mean of the observed data.

PBIAS is a measure of the model’s tendency to either over or under-predict, and is calculated as follows:

If the model on average over predicts, PBIAS is greater than 0. Under predictions result in negative PBIAS values.

The third metric is designed to give a description of the model’s absolute error, and is calculated as follows:

It is suggested that .

# 4 Model Calibration and Validation

## 4.1 Sensitivity Analysis

In order to calibrate a model, it is important to use only those model parameters that have a strong influence on model output. There are two ways to determine which parameters are sensitive. A “global” sensitivity analysis can be done by varying multiple parameters concurrently and computing a multiple linear regression function between the parameter and objective function values. While this captures the full variability of each parameter, the interactions between parameters can obscure the results. The second way is called “one-at-a-time” sensitivity analysis and involves changing one parameter at a time and measuring the associated change in model output. This isolates the effect of each parameter but misses information on the interaction between multiple parameters. This limitation can be alleviated with a proper understanding of the model algorithms, but some feedbacks throughout the system will still be missed. Since calibrating for flow, sediment and nutrients involves a large set of parameters, we chose to conduct “one-at-a-time” sensitivity analysis. We used the SWAT-CUP calibration software (Abbaspour 2012) which provides detailed representations of sensitivity results. Figure 3 displays an example of this process showing results for the parameter EPCO (Plant uptake compensation factor). As the parameter increases, winter flow decreases. Sensitivity analysis was done for 43 parameters. A summary of sensitive parameters can be found in Table 4.

Fig 4**:** Sensitivity analysis of the EPCO parameter (Plant uptake compensation factor) using the SWAT-CUP software. Larger values of EPCO result in smaller peak flows during winter months.

## 4.2 Calibration and Validation

There are two main strategies for determining model parameters: Manual and automated calibration. A modeler using the manual approach will run the model, and then adjust parameters individually until a good fit to the observed data has been made. Automated calibration routines vary the model parameters systematically over the course of many model simulation runs and use objective (numeric) criteria to determine which runs are most accurate. The calibration results for automatic calibrations are sensitive to the objective criteria chosen. Krause et al. (2003) discuss the common objective criteria used, and how each impacts calibration results. Whichever method is used, the final parameter set should be representative of the physical characteristics of the basin. However, due to the spatial and temporal variation in these characteristics, as well as the difficulty in acquiring direct measurements of these parameters, it is difficult to know what the true parameters are (Abbaspour et al 1999). This results in a large degree of uncertainty in parameter estimates. Furthermore, there are often times more than one set of parameters that provide a good model fit. This is referred to as the equifinality, or non-uniqueness problem (Abbaspour 2012).

We used data on our study site to perform manual calibration. Three gages were used to calibrate the Tualatin and one gage was used to calibrate the Yamhill. Parameters selected were all global in nature with the exception of in-stream sediment deposition and erosion parameters which were localized to the three gages in the Tualatin in instances where the MUSLE alone did not provide adequate model fit (Table 5).

## 4.4 Calibration and Validation Results

### 4.4.1 Flow

Based on a review of the existing literature and our sensitivity analysis, six parameters were used to calibrate the Tualatin. Five were used to calibrate Yamhill. Monthly calibration results for flow were quite good in both basins, with NSE ranging between 0.85-0.94 during the calibration period and 0.7 – 0.9 for the validation period (Figure 3). The simulated runoff ratio for Tualatin is 0.58 and 0.65 for Yamhill (Table 4). Despite the heterogeneous land cover throughout the Tualatin basin, a global parameter set was sufficient to acquire a good flow calibration at each of the three gages.

The SWAT models underestimate flow consistently throughout both basins for both calibration and validation periods. In the Tualatin, the most severe underestimation was in Fanno creek during both calibration and validation periods (Fig 3). The McMinville gage had the most severe underestimation of all gages in the validation period (-37.7%). West Linn had the smallest bias (< 2%) in both calibration and validation periods. Visual comparison of the observed and simulated results by hydrograph show that the source of bias in the model is mostly wintertime high flows, however there are some instances of underestimation during low flow periods as well in calibration and validation periods, suggesting that evapo-transpiration is too high. RSR values never exceed 0.54 at all gages for both calibration and validation periods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 4**: Simulated historic water budget (mm) for the time period 1/1/1979 - 12/31/2005 | | | | | |
| **Basin** | **Precip** | **Runoff** | **PET** | **AET** | **Runoff Ratio** |
| ***Tualatin*** | 1,212 | 709 | 939 | 488 | 0.58 |
| ***Yamhill*** | 1,389 | 900 | 967 | 467 | 0.65 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table 5:** List of final calibrated parameters for Tualatin and Yamhill sub-basins | | | | | | |
| **Description** | **Parameter** | **Min** | **Max** | **Sensitivity (Increase/Decrease)** | **Tualatin Fitted Value** | **Yamhill Fitted Value** |
| ***Flow*** |  |  |  |  |  |  |
| SCS runoff curve number | r\_\_CN2.mgt\* | -1 | 1 | I | 0.2 | 0.2 |
| Baseflow alpha factor (days) | v\_\_ALPHA\_BF.gw\* | 0 | 1 | Rising limb: I | 1 | 1 |
| Falling limb: D |
| Soil evaporation compensation factor | v\_\_ESCO.bsn | 0.01 | 1 | I | 0 | 0 |
| Plant uptake compensation factor | v\_\_EPCO.bsn | 0 | 1 | D | 1 | 1 |
| Available water capacity of the soil layer | r\_\_SOL\_AWC().sol | -0.2 | 0.2 | D | -0.2 | -0.2 |
| Treshold depth of water in the shallow aquifer required for return flow to occur (mm) | v\_\_GWQMN.gw | 0 | 5000 | D | 0.1 | 0 |
| ***Sediment*** |  |  |  |  |  |  |
| USLE equation support practice factor | r\_\_USLE\_P.mgt | -1 | 1 | I | -0.6 | 0 |
| Average slope length | r\_\_SLSUBBSN.hru | -1 | 1 | I | -0.3 | 0.25 |
| Min value of USLE C factor applicable to the land cover/plant (Forest) | r\_\_USLE\_C.crop.dat | -1 | 1 | I | -0.3† | 0.35† |
| -0.2†† | 0.25†† |
| -0.2††† | 0.25††† |
| USLE equation soil erodibility (K) factor | r\_\_USLE\_K().sol | -0.2 | 0.5 | I | -0.5 | 0 |
| Channel erodibility factor | v\_\_CH\_COV1.rte | 0 | 1 | I | 0.3‡ | 0 |
| 0.9‡‡ |
| 0.1‡‡‡ |
| Channel cover factor | v\_\_CH\_COV2.rte | 0 | 1 | I | 0.3‡ | 0 |
| 0.9‡‡ |
| 0.1‡‡‡ |
| Linear parameter for calculating the maximum amount of sediment that can be reentrained during channel sediment routing | v\_\_SPCON.bsn | 0.0001 | 0.01 | I | 0.0016 | 0.0001 |
| ***Nitrogen*** |  |  |  |  |  |  |
| Concentration of nitrogen in rainfall | v\_\_RCN.bsn | 0 | 10 | I | 1 | 1 |
| Nitrogen percolation coefficient | v\_\_NPERCO.bsn | 0.1 | 1 | I | 0.01 | 0.01 |
| Organic N enrichment ratio | v\_\_ERORGN.hru | 0.1 | 5 | I | 2.5 | 0.1 |
| Denitrification exponential rate coefficient | v\_\_CDN.bsn | 0.1 | 3 | I | 0.2 | 0.2 |
| Denitrification threshold water content | v\_\_SDNCO.bsn | 0.1 | 1 | D | 1 | 1 |
| ***Phosphorus*** |  |  |  |  |  |  |
| Initial organic P concentration in surface soil layer | v\_\_SOL\_ORGP().chm | 0 | 100 | D | 0.01 | 0.05 |
| Initial labile (soluble) P concentration in surface soil layer | v\_\_SOL\_SOLP().chm | 0 | 100 | I | 1 | 0.05 |
| Phosphorus soil partitioning coefficient | v\_\_PHOSKD.bsn | 100 | 200 | I | 150 | 200 |
| Phosphorus Enrichment Ratio | v\_\_ERORGP.hru | 0 | 5 |  |  |  |
| \*v: Parameter is assigned this value. r: Parameter is increased or decreased by this multiple. | | | | | | |
| †Evergreen(6), deciduous(7), mixed(8), wetland(10) forest †† Row crops (2) ††† Rangeland (15, 16) | | | | | | |
| ‡Reaches upstream of Dilly ‡‡Reaches upstream of Fanno ‡‡‡All other reaches | | | | | | |

### 4.4.2 Sediment

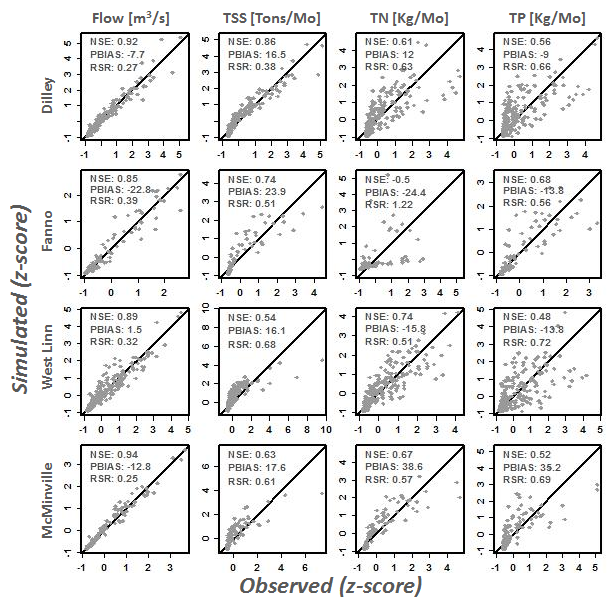
Sediment results were mixed. NSE values ranged from 0.54 – 0.86 during calibration and 0.48 – 0.80 during validation. West Linn had the lowest NSE scores in both periods while Dilley had the best. The model consistently overestimates sediment, with bias ranging from 16.1 – 23.9% during calibration and 7.1 – 102.4% during validation. While over-all model bias is positive, it is clear from the time series (Table 5 and 6) that sediment is underestimated during high flows and overestimated during medium and low flows. If high flow years of WY 1996 – 1999 are removed, the bias for sediment at West Linn increases to 62.3%. RSR ranges from 0.38 (West Linn) – 0.68 (Dilley) during calibration and 0.44 (West Linn) – 0.71 (Dilley).

### 4.4.3 Total Nitrogen

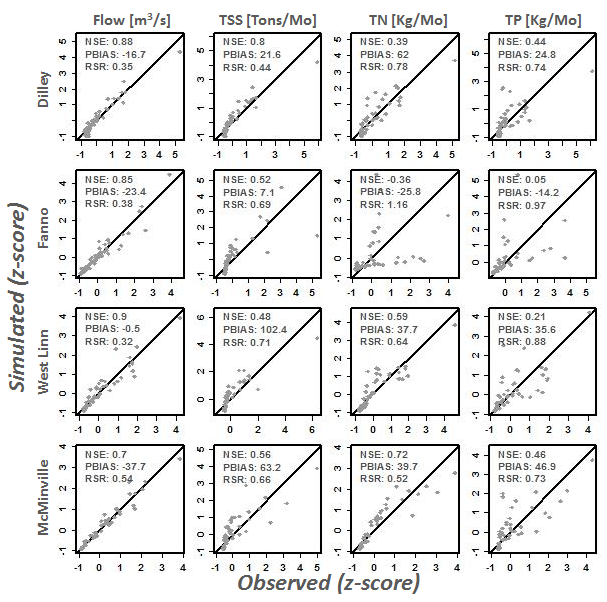
NSE scores range from -0.5 – 0.74 during calibration and -0.36 – 0.72 during validation. Fanno is the only basin with a negative NSE score during calibration and validation. Bias ranges from -24.4% – 38.6% during calibration and -25% – 62% during validation. These values are large, but still within the acceptable range according to Moriasi et al (2007). RSR values range from 0.51 – 1.22 during calibration and 0.52 – 1.16 during validation. Fanno is the only gage which has an RSR value above 0.7 during calibration. Both Fanno and Dilley have RSR values above 0.7 during validation. Based on these results, the areas upstream of Fanno Creek, the most urbanized watershed, do not properly model TN processes.

### 4.4.4 Total Phosphorus

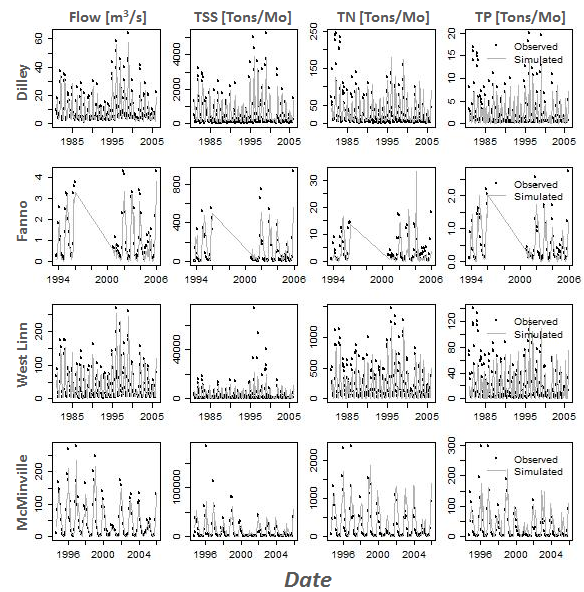
NSE scores ranged from 0.48 – 0.68 for calibration and 0.05 – 0.46 for validation. All NSE scores were worse during validation. Fanno dropped the most, from 0.68 to 0.05, while McMinville dropped the least, from 0.52 to 0.46. Bias ranged from -13.8 – 35.2% for calibration and -14.2 – 46.9% during validation. These values were not consistent, however. As can be seen in Figures 3 and 4, the models had a wide range of both under and over estimation. RSR values range from 0.56 – 0.72 for calibration and 0.73 – 0.97 for validation. West Linn was the only gage to exceed 0.7 during calibration while all gages exceeded this value during validation.



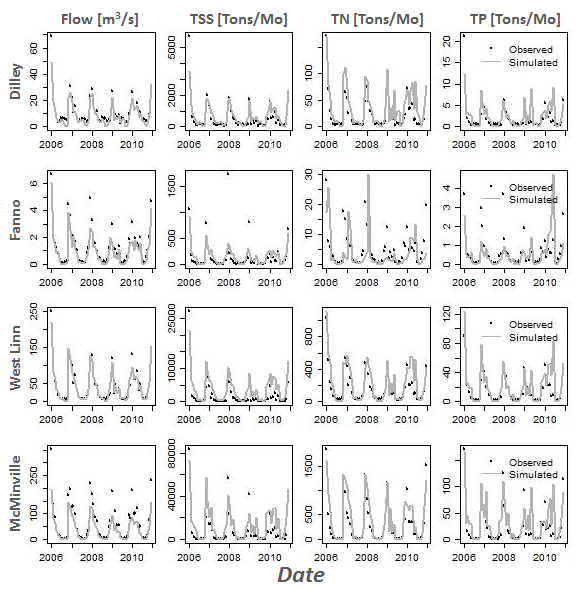
**Fig. 5:** Monthly calibrated results for both the Tualatin and Yamhill sub-basins.



**Fig. 6:** Validation results for Tualatin and Yamhill sub-basins.



**Fig. 7:** Time series of calibrated results for both the Tualatin and Yamhill sub-basins.



**Fig. 8:** Time series of validated results for both the Tualatin and Yamhill sub-basins.

## 4.5 Spatial patterns of flow, nutrients, and sediment

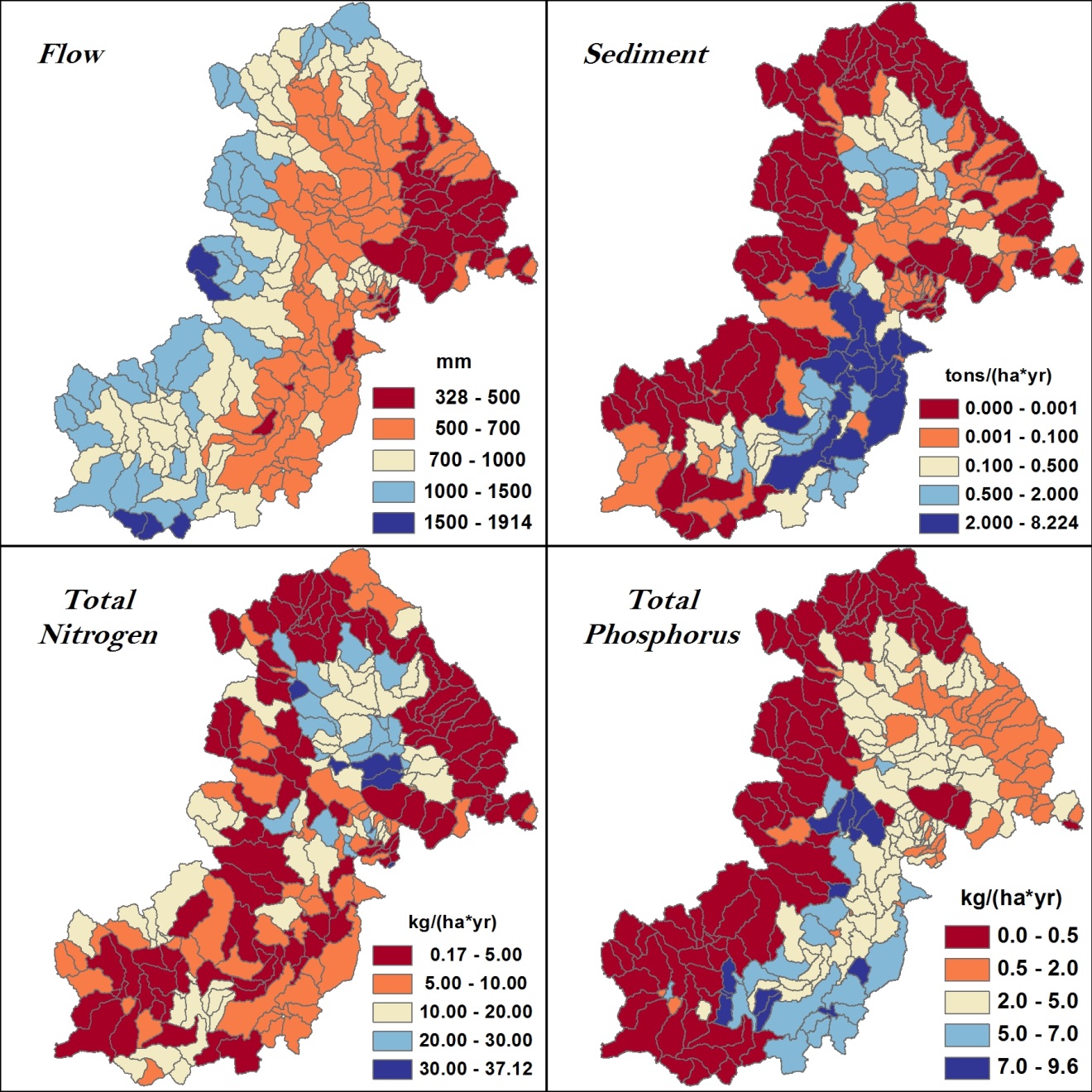
The spatial patterns of flow, sediment and nutrients can be seen in Figure 7. Spatial patterns for water yield are ultimately determined by topography. Orographic effects from the Coast Range cause higher precipitation rates at higher elevations resulting in a clear east-west gradient with smaller yields in the east, and larger yields in the west.

There are two distinct patterns in sediment yields for the basins. Yields in the Tualatin are all uniformly small. This is most likely due to misallocation of sediment sources and sinks in the model. Evaluation of sediment yields from the modeled reservoir after calibration revealed that sediment trapping is underestimated. Since the sediment MUSLE parameters for the Tualatin model were calibrated to the Dilley gage first, the sediment yields from the landscape had to be reduced to compensate for the artificially high levels of sediment passing through the dam, resulting in artificially low yields from the landscape. Furthermore, the Tualatin River has a check dam at RM 3.1, upstream of the West Linn gage, which is not included in the model. While this dam did not affect the MUSLE parameterization at the Dilley gage, it none-the-less is an additional sediment trap that is not properly accounted for in the model.

In contrast, yields in the Yamhill show a more heterogeneous spatial pattern. Areas with heavy agriculture have higher yields than forested areas. Upland areas still have very low sediment yields. However we have no data available with which to estimate the yields that should be expected from these areas.

Higher total nitrogen yields are seen on agricultural lands in the Tualatin basin while urban and forested lands have significantly less yield. In the Yamhill, the relationship to land use is less apparent. Agricultural sub-basins have lower nitrogen yields than some forested higher elevation sub-basins. This is may be due to the organic nitrogen enrichment ratio parameterized lower in the Yamhill than in the Tualatin, resulting in lower levels of organic nitrogen being transported from the surface to streams along with colloidal sediment particles.

Soluble phosphorus particles are transported with sediment particles along with surface runoff, and this pattern can be seen in the results for Yamhill where sub-basins with high sediment yields also have high phosphorus yields. Whether this pattern holds for Tualatin is difficult to determine with the indistinguishably small sediment yields. However, the higher levels of phosphorus originating from the agriculture dominated sub-basins means that the spatial patterns make intuitive sense.



**Figure 9:** Spatial patterns of flow, sediment, total nitrogen, and total phosphorus.

# 5 Discussion

## 5.1 Sources of Uncertainty

### 5.1.1 LOADEST

These models are “roughly” calibrated and have a variety of sources of uncertainty which should be mentioned. First, while LOADEST is commonly used in SWAT studies (Almendinger & Murphy 2007, Kannan 2012), there are uncertainties in any load estimates. Flow values are used to calculate loads from concentration estimates made from grab samples. Thus, while correlations between concentration and flow are generally small, high correlations between flow and loads exist due to the fact that flow is used to compute both the independent and dependent variables. Shivers and Moglen (2008) refer to this as spurious correlation. LOADEST does provide 95% confidence intervals for all outputs, but these are for the calibrated dataset, not the estimation dataset. This may not be a problem in situations where there is no extrapolation outside the range of flow values in the calibration dataset. This is not the case for our study, however, and is an acknowledged limitation (Table 6). Days which did not have all constituents comprising TN or TP were excluded from LOADEST models. Since SWAT algorithms model each constituent separately, developing load estimates of nitrate/nitrite, organic nitrogen, ammonia, soluble, and insoluble phosphorus separately may facilitate a more precise calibration of SWAT nutrients. This may also enable comparison with Ammonia and Phosphorus TMDLs.

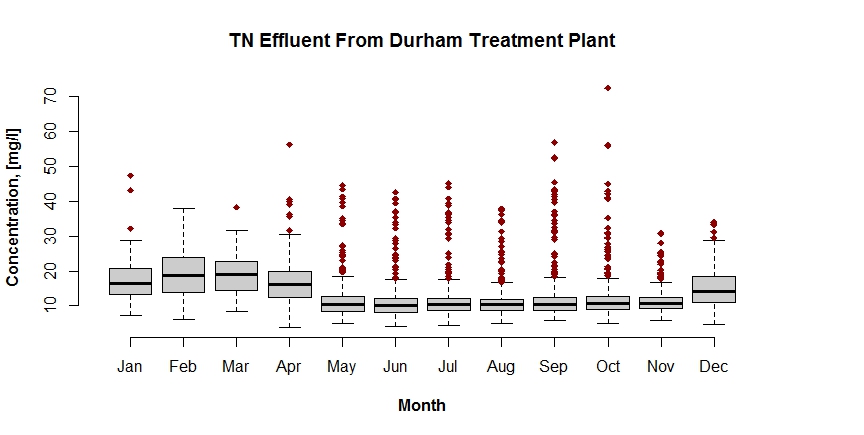
|  |  |  |  |
| --- | --- | --- | --- |
| **Table 6:** Percent difference between maximum flow in the calibration period (CP) and estimation period (EP). | | | |
| ***Gage*** | ***Maximum Flow (cfs) EP*** | ***Maximum Flow (cfs) CP*** | ***Difference (%)*** |
| **Dilley** | 8620 | 3350 | 157.3 |
| **Fanno (TSS)** | 1410 | 779 | 81.0 |
| **Fanno (TN, TP)** | 1410 | 840 | 67.9 |
| **West Linn** | 25900 | 15500 | 67.1 |
| **Yamhill DEQ** | 56420 | 17080 | 230.3 |

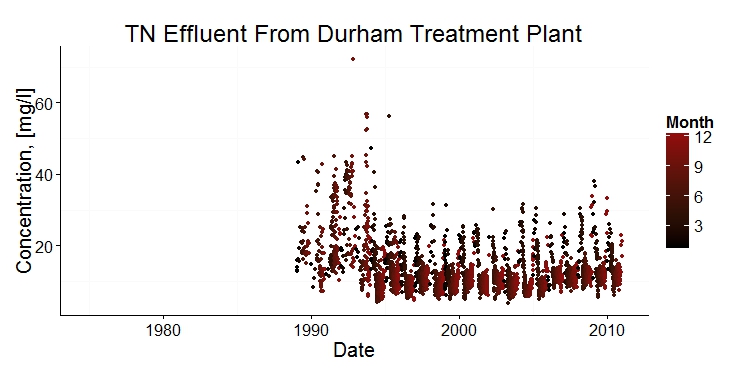
### 5.1.2 Parameter Uncertainty

Another area of uncertainty lies in the measurement and estimation of model parameters. Because it is often times difficult or impossible to measure these parameters it is becoming generally accepted as unrealistic to ascribe one value to all parameters in a watershed model. Rather, information on the study site should be collected so that informed estimates of sensitive parameters can be made. Once realistic parameter ranges have been chosen, parameter uncertainty can be assessed through a Monte Carlo style approach (Abbaspour et al 2007; Almendinger and Murphey 2007) using the SWAT-CUP software. While this type of uncertainty analysis was beyond the scope of this study, future work should consider including this analysis when presenting calibration results.

### 5.1.3 Waste-water Treatment Plants

The presence of WWTPs along the Tualatin River poses a challenge to accurate modeling. While the flow from the WWTPs was included in the model, deriving daily or monthly estimates of nutrient effluent proved difficult due to lack of historic data and implementation of TMDLs in the mid 1990’s which helped reduce nutrient loads (Figure 4), but created distinct changes in temporal loading patterns. Since these loads are not included in the model, nutrients off the landscape are inherently over-estimated.





**NO DATA**

**AVAILABLE**

**Fig 10:** Box plots and time series of Total Nitrogen concentrations of effluent from the Durham waste-water treatment plant.

### 5.1.4 Urban Areas

Urban land uses also present a challenge. Because impervious areas create flashier runoff responses to storms which sometimes take place at sub-daily timescales, current algorithms using the daily timestep are not always reliable. Furthermore, SWAT cannot simulate the first flush effect at the daily timestep, suggesting that the model may overestimate sediment and nutrient loads. While SWAT was developed as a tool for agricultural areas, new routines for urban runoff and sediment transport are under development that may improve results for small urban watersheds (Jeong et al 2010; Jeong et al 2011). In instances when modeling urban and suburban areas is critical, HSPF is a better option, although the user interface is not as intuitive as SWAT’s (Borah & Bera 2003, Borah et al 2006, Saleh and Du 2004).

## 5.4 SWAT Modeling for Ecosystem Services

Despite certain inherent limitations with SWAT, such as event scale, and urban/sub-urban land use modeling, the SWAT models presented here could be used to measure the biophysical output these basins can be expected to produce under a given set of climatological and land use conditions. In economic terms, this is known as the ecological production function (Tallis & Polasky 2009). This type of information is useful for regulatory agencies that are responsible for setting and enforcing water quality standards. Further information on how SWAT can be pared with other models such as CEQUAL-W2 or to compensate for SWAT’s limitations can be found in Borah et al (2006). However, there are limitations to how far the biophysical models by themselves can be employed in these situations. For example, a successful model can be used to estimate changes in timing and amount of flow, sediment and nutrient loads at the calibrated gages. It can also be used to see where potential water quality “hotspots”, also known as critical source areas (CSAs), may be in the basin (Niraula et al 2011, Niraula et al 2013). However, none of these analyses include economic tradeoffs. This second level of ecosystem service analysis involves measuring the costs and economic and biophysical tradeoffs associated with water quality improvements. This type of analysis is critical for enabling regulatory agencies to make informed land management decisions.

Rabotyagov et al (2010) uses estimated costs for grassed waterways, land rental rates, costs for no-till farming practices and terraces, and a genetic algorithm with SWAT to determine optimum land use scenarios for reaching a 30% reduction in Nitrate-N or TP in the Upper Mississippi River Basin, USA. They found that optimizing for Nitrate-N had an associated reduction in TP of 36% and cost $1.4 billion/yr, while optimizing for TP had an associated reduction in Nitrate of 9% and cost $370 million/yr. Bekele and Nicklow (2005) demonstrated the effectiveness of using an integrative modeling framework that pairs SWAT with the multiobjective evolutionary algorithm SPEA2 to evaluate tradeoffs. Bekele et al (2013) used this modeling framework with cost data to develop ecological-economic production possibility frontiers for the Big Creek watershed in southern Illinois. They found that while in aggregate provisioning services (crops) and regulating services (flood protection, water quality, and carbon retention) have an inverse relationship, a small reduction in provisioning services brings a large increase in regulating services.

While using SWAT to assess economic tradeoffs is a relatively new field, and has inherent limitations to the types of ecosystem services it can measure, its ability to optimize solutions is promising. Limitations to widespread use of these methods must be overcome, however. Current studies linking SWAT to genetic algorithms require familiarity with code development since there is no software immediately available to do this as far as the authors of this paper are aware. However, as the SWAT-CUP calibration and uncertainty analysis software shows, it is possible to design software that seamlessly integrates with the SWAT model. Future software development could involve linking SPEA2 to SWAT using a graphical user interface so that watershed modelers and environmental economists can more easily work together to measure ecosystem service tradeoffs.

# 6 Conclusions

As land use and water quality managers begin using the ecosystem service framework to make decisions, assessments of the available ecosystem service tools must be made. In this article we calibrate a popular ecosystem service model, SWAT, for two sub-basins in Northwest Oregon and discuss the uncertainties and issues involved. We found that the model has a robust suite of tools available to help with input data, calibration and uncertainty analysis. The LOADEST software helps simulate sediment and nutrient loads in locations where continuous sampling is not available. The SWAT-CUP software helps with sensitivity analysis, calibration, and uncertainty analysis. While uncertainty analysis historically has not been quantified in watershed modeling research and we were unable to perform it in this project, the SWAT-CUP software makes the process simpler, and we suggest future SWAT models include this in their analysis.

Results of model calibrations were mixed. Flow simulations track well with observed data in both basins, and the spatial patterns of water yield make sense given the local topography. Sediment simulation results in the Tualatin were acceptable at the Dilley and Fanno Creek gage, but the West Linn gage is in need of improvement. Sediment results for the Yamhill DEQ gage were also acceptable. Spatial patterns of sediment yield are sensible for the lowland areas of Yamhill, but may be underestimated in the upland regions. Sediment yield in the Tualatin register below SWAT’s reporting level in many cases. This is likely due to high levels of sediment passing through the modeled Hagg Lake reservoir, causing sediment to be underestimated across the landscape. Total Nitrogen calibrations are acceptable for the West Linn and Yamhill DEQ gage, but were unsatisfactory for the Dilley and Fanno gages. Spatial patterns of TN yield appear sensible in the Tualatin where yields are higher in agricultural lands, while there is no distinct pattern in Yamhill. This may be due to underestimation of organic nitrogen in the Yamhill basin, and suggests the SWAT model should be calibrated to each species of nutrient to ensure proper budgeting throughout the system. Finally, Total Phosphorus calibration results were poor in all cases. Spatial patterns of TP yields are sensible in both basins with higher yields in agricultural lands.

While the models calibrated in this study can in some cases be used to evaluate the ecological production function, further economic analysis requires additional tools. Examples include the inclusion of multi-objective evolutionary algorithms to optimize reductions in water pollution, and evaluate economic tradeoffs through the development of ecological-economic production possibility frontiers. While these tools show promise, new user-friendly software linking SWAT to these algorithms need to be developed before widespread use of these methods can be achieved.

The results of this study suggest that SWAT is a useful tool for the evaluation of ecosystem services, and that future developments in the field may facilitate its continued widespread use for both biophysical analysis and economic analysis of ecosystem services.

# References

Abbaspour KC, Sonnleitner MA, Schulin R (1999) Uncertainty in estimation of soil hydraulic parameters by inverse modeling: Example lysimeter experiments. Soil Sci Soc Am J 63:501–509.

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

Abbaspour KC (2012) SWAT-CUP 2012: SWAT Calibration and uncertainty programs – A user manual. EAWAG.

Abatzoglou JT (2013) Development of gridded surface meteorological data for ecological applications and modeling. Int J Climatol 33:121­–131.

Almendinger JE, Murphey MS (2007) Constructing a SWAT model of the Willow River watershed, western Wisconsin. St. Croix Watershed Research Station, Science Museum of Minnesota: 1-84.

Arnold JG, Moriasi DN, Gassman PW, Abbaspour KC, White MJ, Srinivasan R, Santhi C, Harmel RD, van Griensven A, Van Liew MW, Kannan N, Jha MK (2012) SWAT: Model use, calibration, and validation. Trans of ASABE 55(4):1491-1508.

Bekele, EG, Lant CL, Soman S, Misgna G (2013) The evolution and empirical estimation of ecological-economic production possibilities frontiers. Ecol Econ 90:1-9.

Boeder, M., H. Chang. 2008. Multi-scale analysis of oxygen demand trend in an urbanizing Oregon watershed. *Journal of Environmental Management* 400(1-3): 567-581.

Brauman KA, Daily GC, Duarte TK, Mooney HA (2007) The nature and value of ecosystem services: An overview highlighting hydrologic services. Annu Rev Environ Resour 32:6.1-6.32.

Brown LC, Barnwell Jr TO (1987) The enhanced water quality models QUAL2E and QUAL2E-UNCAS documentation and user manual. USEPA, Athens, GA.

Chang, H. and J. Jones. 2010. Climate change and freshwater resources in Oregon. In Oregon *Climate Assessment Report*, eds. K.D. Dello and P.W. Mote. 69-150 College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR.

Cochran B, Logue C (2011) A watershed approach to inprove water quality: Case study of Clean Water Services’ Tualatin river program. JAWRA 47(1):29–38.

Costanza, R., R. d’Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R.V. O’Neill, J. Paruelo, R.G. Raskin, P. Sutton, and M. van den Belt. 1997. The value of the world’s ecosystem services and natural capital. *Nature* 387: 253-260.

Fry JA, Xian G, Jin S, D JA, Homer CG, Yang L, Barnes CA, Herold ND, Wickham JD (2011) National land cover database for the conterminous United States. *PE&RS 77*(9):859–864.

Franczyk, J., H. Chang. 2009. The effects of climate change and urbanization on the runoff of the Rock Creek basin in the Portland metropolitan area, Oregon USA. *Hydrol. Process* 23: 805-815.

Garen, D.C., D.S. Moore. 2005. Curve number hydrology in water quality modeling: Uses, abuses, and future directions. *JAWRA* 41 (2): 337-388.

Green WH, Ampt GA (1911) Studies on soil physics, 1. The flow of air and water through soils. J Agri Sci 4:11-24.

Hamlet, A.F., D.P. Lettenmaier. 1999. Effects of climate change on hydrology and water resources in the Columbia River Basin. *JAWRA* 35 (6): 1597-1623.

Hernandez M, Kepner WG, Goodrich DC, Semmens DJ (2010) The use of scenario analysis to assess water ecosystem services in response to future land use change in the willamette river basin, Oregon. In: Liotta W, Kepner JL, Mouat D (eds) Achieving environmental security: Ecosystem services and human welfare. IOS Press, Amsterdam, Netherlands, 69:265-268.

Kannan, N. 2012. SWAT modeling of the Arroyo Colorado Watershed. Technical Report No. 4262. Texas Water Resources Institute, Temple, TX.

Kremen, C., R.S. Ostfeld. 2005. A call to ecologists: Measuring, analyzing, and managing ecosystem services. Front Ecol Environ 3(10): 540-548.

Liang X, Lettenmaier DP, Wood EF, Burges SJ (1994) A simple hydrologically based model of land surface water and energy fluxes for GSMs. Geophys Res 99(D7):14,415-14,428.

Logsdon, R.A., I. Chaubey. 2013. A quantitative approach to evaluating ecosystem services. *Ecological Modelling* 257: 57-65.

Millennium Ecosystem Assessment Report (MA) (2005) Ecosystems and human well-being: Wetlands and water synthesis. World Resources Institute, Washington DC.

Moriasi, DN, Arnold JG, Van LIew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans of the ASABE 50(3):885-900.

Moriasi, D.N., J.L. Steiner, J.G. Arnold. 2011. Sediment measurement and transport modeling: Impact of riparian and filter strip buffers. *J. Environ. Qual.* 40: 807-814.

Mukundan R, Radcliffe DE, Risse LM (2010) Spatial resolution of soil data and channel erosion effects on SWAT model predictions of flow and sediment. Jour Sol Wat Cons 65(2):92–104

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, J.R. Williams. 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009. Texas Water Resources Institute Technical Report No. 406.

NHD Plus 2006. NHDPlus User Guide <http://www.fws.gov/r5gomp/gom/nhd-gom/NHDPLUS_UserGuide.pdf>. Last accessed 3/1/2013.

Niraula R, Kalin L, Srivastava P, Anderson CJ (2013) Identifying critical source areas of nonpoint source pollution with SWAT and GWLF. Ecol Modelling 268:123-133.

Norman LM, Villarreal ML, Niraula R, Meixner T, Frisvold G, Labiosa W (2013) Framing scenarios of binational water policy with a tool to visualize, quantify and valuate changes in ecosystem services. Water 5:582-874.

Oregon Department of Environmental Quality (ODEQ) 2001. Tualatin Subbasin Total maximum Daily Load (TMDL): <http://www.deq.state.or.us/wq/tmdls/docs/willamettebasin/tualatin/tmdlwqmp.pdf>. Last accessed 3/1/2013.

Praskievicz, S. and H. Chang. 2011. Impacts of climate change and urban development on water resources in the Tualatin river basin, Oregon. *Annals of the Association of American Geographers* 10 (2): 249-271.

Rabotyagov, S, Campbell T, Jha M, Gassman PW, Arnold J, Kurkalova L, Secchi S, Feng H, Kling CL (2010) Least-cost control of agricultural nutrient contributions to the Gulf of Mexico hypoxic zone. Ecol App 20(6):1542-1555.

Runkel RL, Crawford CG, Cohn TA (2004) Load estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. Techniques and Methods Book 4, Chapter A5, USGS, Reston, VA.

Shivers DE, Moglen GE (2008) Spurious correlation in the USEPA rating curve method for estimating pollutant loads. J Environ Eng 134:610–618.

Sobieszczyk, S (2012) Annotated Bibliography – Tualatin river basin studies. USGS. <http://or.water.usgs.gov/tualatin/>

Soil Conservation Service (1972) Section 4: Hydrology In National Engineering Handbook. SCS.

Tallis P, Polasky S (2009) Mapping and valueing ecosystem services as an approach for conservation and natural-resource management. Ann NY Acad Sci 1162:265-283.

Quintero M, Wunder S, Estrada RD (2009) For services rendered? Modeling hydrology and livelihoods in Andean payments for environmental services schemes. Forest Ecol Manage: 1-10.

Vigerstol, K.L., J.E. Aukema. 2011. A comparison of tools for modeling freshwater ecosystem services. *Journal of Environmental Management* 92(2011): 2403-2409.

Williams J (1975) Sediment routing for agricultural watersheds. Wat Res Bul 11(5):965–974.