6. LIMITATIONS OF THE STUDY

Any modeling exercise is an attempt to simulate real world processes through the use of input data describing physical characteristics of the system, a set of algorithms to transform input data to output parameters of interest, and simplifying assumptions to limit the scope of the model. The following aspects of the study imposed limits on the accuracy of the simulated output and comparisons with the monitored data.

6.1 Data Limitations

6.1.1 Monitored Data

The OWML monitoring data set was provided as a working document, unchecked for errors. While a variety of error checks were run on this data base, and flow records checked with a parallel USGS gauge for part of the time period, assumptions still had to be made, and personal judgment used on the appropriateness of certain data records. Also, where interim baseflow records were missing, default values were used to estimate flow so that monthly flow could be calculated from the monitored data.

6.1.2 Rainfall

Rainfall data were assessed at the three closest rain gauges to the watershed, and The Plains station chosen as it showed the greatest influence from a Thiessen weighting procedure, in agreement with the best alignment of rainfall with composite-sampled time intervals. Considerable variation in rainfall existed amongst these three stations, and suggested that rainfall distribution was not uniform over the watershed, as well. Research shows that rainfall variability within a watershed can be considerable, and is also affected by the direction and intensity of the storms (Goodrich et al., 1993; Troutman, 1983; Young et al., 1992; DiLuzio and Lenz1, 1995). A comparison of rainfall monitored at The Plains station with runoff monitored at the Bull Run ST60 station revealed an unnaturally high percentage of runoff for many of the storms indicating a poor match between rainfall and runoff. Since The Plains rainfall was used to produce modeled runoff and loads, but does not represent the rainfall which produced the monitored runoff and loads, the comparison between modeled and monitored runoff and loads contained many uncertainties. This, unfortunately, invalidated the comparison and rendered the evaluation of model performance inconclusive.

6.1.3 Land Use

The annual land use data matrix was not comprised of actual land uses, but was interpolated, deduced, and extrapolated from limited category land use scenes for 1981 and 1990. Disturbed areas were not identified as they were mostly transient within any given year, though they could be responsible for considerably larger amounts of runoff and loading than with other land uses with greater ground cover. Because actual crop histories were not available on a field-by-field basis, four predominant crop rotations were used to assign agricultural land use on a year-by-year and cell-by-cell basis. Variations from the predominant crop rotations, such as turf and sod farming, and occasional livestock in the watershed were not represented in the land use characteristics or modeling options used. Furthermore, one land use category, that of unspecified agriculture, was somewhat ambiguous. This category may have included farmsteads, lawns, disturbed areas, idle land, and other land uses not identifiable as one of the other specified land use categories. Highways and roads were not modeled uniquely, but were combined within the rural residential category, reducing the larger amounts of runoff and different types of loading normally expected from these impervious areas.

6.1.4 Daily Erosion Index (EI)

The energy associated with each storm was not calculated from breakpoint data as recommended by Wischmeier and Smith (1978), but was estimated from a regression equation with monthly coefficients, developed for another region of the state. Even if the month-to-month distributions for Bull Run were similar to the site for which the regression was developed, which is itself questionable, large differences could still have occurred between actual and estimated EI on an event basis, especially with smaller storms.

6.1.5 Resolution of Spatial GIS Data

Values assigned to any raster or grid cell represents an average value over the area of each cell. The greater the variability over the cell, the greater will be the error induced through the use of an average value. The data in this study were generally aggregated from 1/9 ha to 4/9 ha grid cells to accommodate the total number of cells constraint within the AGNPS 5.0 model. Because of this aggregation, the spatial variability of values within any data layer were compressed and the value in any given cell represents an average from a larger range of values within the larger cell.

6.2 Model Limitations

6.2.1 Event-Based Model

AGNPS is an event-based model which simulates surface runoff and does not have a groundwater component. Therefore, any comparison with monitored data must either subtract groundwater from the monitored data, or add groundwater to the modeled data for realistic evaluations of model performance.

6.2.2 Known Bugs

AGNPS 5.0 has a number of known bugs, acknowledged by the developers, but currently not fixed. The one most applicable to this study is in the development of the hydrograph steps, where it has been shown to overestimate sediment loads by 10%.

6.2.3 Curve Number Basis

AGNPS generates runoff from an empirical-based equation rather than from equations more closely representing physical processes. Though definitely a limitation, the use of expanded lists of CN by NRI crop-tillage-management combinations and the use of seasonally-variable CN for cropland, should have improved its application over the standard CN approach.

6.2.4 TR-55

The inability of the TR-55/geomorphic combination to model small storms accurately limited the range and number of storms in the comparison to larger storms, and eliminated comparisons on a monthly basis for months where only smaller storms occurred within a month.

6.2.5 NonFeedlot Point Source Option

Impervious areas were not modeled using buildup/washoff functions as intended due to an error in the AGNPS algorithms. This reduced the contributions of these areas from both runoff and loading.

6.2.6 Default Parameters

AGNPS employs many parameters which are used over the entire watershed. As was shown in the results, total phosphorus generally overpredicted monitored levels especially for smaller storms. Most likely, this resulted from inappropriate default phosphorus values for potency factors for soil or water, partitioning coefficients for leaching and runoff, or decay rates within stream flow.

6.3 Assumption Limitations

Many assumptions were used in the monthly simulation modeling of baseflow and septic system loads. Monthly baseflow was assumed to equal the average of the monthly mean and minimum daily flow times the number of days in a month. Concentrations of N in the baseflow were based on national median values by HSG, while a constant median P concentration was used.

In the septic system load modeling, monthly groundwater discharge distribution was assumed equal to monthly runoff distribution. Population was estimated as 3 people per septic system in lieu of exact population distribution figures, and septic systems were considered normal in the absence of information to the contrary.