Data and Variables

Data were chosen that provided the best available regionally consistent spatial coverage for both current and future scenarios. In some cases, better local data were available, but not used because it was not consistently available across the region and an acceptable regional data set was used in its place (e.g., TIGER road data, NLCD land cover). These data represented the important known stressors and resources in the region. All data were reported by 8-digit hydrologic unit codes (HUCs) as they were the only regionally consistent watershed delineation available at the time of this study (Figure 1).

A total of 24 variables was used to relate various environmental, social, and economic dimensions over the 147 8-digit HUCs (Table 1). All variables analyzed were normalized and inverted if necessary to make all indicators range from 0 to 1, where zero and one represent environmentally desirable (low stressor coverage, extensive resource coverage) and undesirable (extensive stressor coverage, low resource coverage) conditions, respectively. Normalization provides a transformation that preserves the ranking and correlation structure of the variables, and allows for variables with different scales to be used together (Pielou 1984).

Landscape Metrics

Land cover and land use variables (percent forest, road density, roads crossing streams, human use and wetland) were generated using the Analytical Tools Interface for Landscape Assessments (ATtILA), version 3.0, an ArcView 3.x extension written by the U.S. EPA (2004). Metrics were based on 1992 National Land Cover Data (NLCD) (projected to 2020 for future scenario), TIGER roads, and National Hydrology Data (NHD) streams to calculate the density, counts, or proportions per HUC. The proportions were determined by summing the variable area and dividing by the total area inside the HUC minus water. Density was calculated as total length divided by area of HUC and counts were summed per HUC. Roads were not weighted for width or lane number.

Forest edge (EDGE2 and EDGE65) and interior (INT2 and INT65) were calculated by moving a fixed sized window across the NLCD land cover at two scales, fine (5 by 5 pixel window covering 2 ha) and coarse (27 by 27 pixel window covering 65 ha) (see Riitters *et al.* 2002). When the center pixel was forest, the number of forest pixels in the window was summed. If the amount of forest in the window was greater than 60%, but less than 100%, the center cell was labeled edge; when all pixels in the window were forest, the center cell was labeled interior.

Forest defoliation was determined using a geographic information system (GIS), to assemble, collate, and analyze gypsy moth defoliation data (Eastman 1989). A 2 x 2 km grid cell size was selected as standard for all map layers in the GIS. The grid size was selected because it represented the minimum dependable spatial resolution of the defoliation data available from state agencies. Data from the suitable habitat combined with forest density, and adjusted for preferred tree species basal area and the predicted geographic pattern of defoliation, were used to predict future potential for gypsy moth defoliation.

Native and Nonnative Species

Data from NatureServe were aggregated by native and nonnative taxa to 8-digit HUCs (see metadata for details). Terrestrial and aquatic species data were based on museum records, literature, expert opinion, and digital databases. A species was considered to be introduced if it did not historically occur in the

Table 1. List of variables (^{S} Stressors and ^{R} Resources) used in scenario analysis.

Abbreviation	Description
AGSL^{S}	Proportion of watershed with agriculture land cover on slopes >9%
AQUAEXOTIC^{S}	Counts of exotic fish and mussel species from heritage data
DISSOLVEDP^{S}	Estimated dissolved phosphorous in streams modeled using land-cover metrics reported as kg/ha/yr
EDGE2^{S}	Percentage of watershed area with forest edge habitat (2 ha scale)
EDGE65^{S}	Percentage of watershed area with forest edge habitat (65 ha scale)
FORCOVDEFOL^{S}	Percent of forest cover defoliated and with mortality as proportion of existing forest
FORPCT^{R}	Percentage of forest coverage
NITRATEGW^{S}	Nitrate in groundwater; probability of exceeding threshold of 3 mg/L
INT2^{R}	Percentage of watershed with interior forest habitat (2 ha scale)
INT65^{R}	Percentage of watershed with interior forest habitat (65 ha scale)
MIGSCENARIO^{R}	Migratory scenarios for long-distance forest migrants
POPDENS^{S}	Population density
POPGROWTH^{S}	Population growth rate from 1990-1995; 2015-2020
PRAGFM^{S}	Percentage of population in agricultural farming
PRMINE^{S}	Percentage of population in mining industry
PSOIL^{S}	Proportion of watershed with potential soil loss greater than one ton per acre per year
RDDENS^{S}	The road density expressed as meters of road per square hectare
RIPAG^{S}	Proportion of total stream length with adjacent agriculture land cover
STRFOR^{R}	Proportion of total stream length with adjacent forest land cover
STRD^{S}	Number of road crossings per total stream length
TERREXOTIC^{S}	Count of exotic birds, mammals, butterflies, amphibians, and reptiles
TOTALN^{S}	Estimated total nitrogen in streams modeled using land-cover metrics
UINDEX^{S}	Human use index (proportion of watershed area with agriculture or urban land cover)
WETLNDSPCT^{R}	Percent of area classified as wetlands

target HUC. The data were compiled on a species-by-species basis. Migratory bird stopover habitats were quantified in 10-km radius hexagons (Tankersley 2004) for the entire study area, modeled on forest density, percent agriculture, and road density (based on NLCD). Details of non-indigenous species spread and migratory bird projections are discussed in the Methods section.

Water

U.S. Geological Survey National Water-Quality Assessment Program groundwater data from studies conducted in the Mid-Atlantic region were used in association with geographic data (land cover, geology, soils, and others) to develop logistic-regression equations that use explanatory variables to predict the likelihood of exceeding nitrate concentration thresholds in shallow aquifers (Greene *et al.* 2005).

Excess export of nitrogen and phosphorus to streams was calculated using a statistical model developed by Reckhow *et al.* (1980) which uses the amounts of land cover to estimate loadings. Weightings for each land cover type were measured in kg per hectare per year. For example, each hectare of urban land cover contributes 1.2 kg of phosphorus and 5.5 kg of nitrogen to a watershed. Final loading values were the sum of contribution by all land cover and uses within a HUC. The loadings model was available in ATtILA (USEPA 2004).

Demographics

Population data were drawn from the Woods and Poole county database (Woods and Poole 2002). Percent change in population from 1990-2000 was then calculated by subtracting the 1990 population from the 2000 population and dividing by the 1990 population. Population density was calculated by dividing projected total population in 2020 by county area. The proportion of the population employed in mining and agricultural farming was calculated by dividing selected employment by total employment. Population data were then apportioned to HUCs using a population-weighted method.