CAPS Input Data Layers – November 2011

This section describes the source input data to CAPS, with a brief listing of major errors and limitations and of modifications we made to the data listed for each source. All data are the most recent available as of November 2011. Most of these data are available from MassGIS (http://www.state.ma.us/mgis/massgis.htm).

MassGIS 2005 land use – This is the source for developed land and some natural types in the landcover. Natural communities from land use include forest, powerline shrublands, and open land. We replaced wetlands and open water in this layer with those in DEP wetlands, which generally align and have higher thematic resolution.

- Water-based recreation is a mix of recreational beaches and associated parking lots and nearby forest, marinas, waterfront promenades, and even roads near beach parking lots.
- Transportation is a hodge-podge of freeway verges, train stations, airports, and highway garages.
- Urban public is a mix of urban hospitals, urban parks, office parks, etc.
- Saltwater sandy beach, which is mostly replaced by DEP wetlands (which breaks it into finer classes); will fill with nearest from DEP wetlands.
- Brushland/successional doesn't appear to be consistently and meaningfully different from open land; we lumped the two.
- Pasture and cropland are often confounded (some off this is real fluidity between these two uses, and much is probably mapping error). Additionally, hayfields are generally mapped as pasture, though the two are different ecologically.
- Cranberry bog polygons are much larger than those in DEP wetlands. They include intervening space around bogs: DEP wetlands mapped just the wet part, land use mapped the use. Land use is more up to date than DEP wetlands, so it maps many new bogs. We union the two and treat them as cranberry bog complexes. Additionally, abandoned cranberry bogs in Burrage Pond Wildlife Management Area are mapped as active cranberry bogs; we changed the class to shallow marsh/meadow/fen.
- We merged very low-density residential with low-density residential (this class isn't consistently applied and doesn't seem particularly meaningful).
- Note that unlike previous versions of Land Use, in this version developed areas are mapped fairly tightly to houses and yards, thus, for instance, wooded suburbs are generally mapped as blobs for individual houses rather than as one big polygon. This makes more sense for CAPS.

DEP wetlands – This is the source of our wetland and open water types, as well as some coastal uplands (beaches, dunes, etc.) DEP wetlands were photo-interpreted, and are generally of high quality, though they are somewhat dated. Beaver pond disturbance/succession has introduced many errors, in particular, current shrub swamps are often mapped as forested wetland.

- We split open water into lotic and lentic based on a model that used the shape of polygons to distinguish rivers from lakes and ponds, followed by thorough hand-checking and editing. For post-processing, we further split lentic into lakes and ponds based on the size of the waterbody (ponds are < 5 ha). This was based on a logistic regression of sizes of lakes and ponds in the National Wetlands Inventory, because NWI distinguishes between lakes and ponds, whereas DEP wetlands depict all open water as one class.
- DEP wetlands were combined with Land Use for all analyses, except for Wetland Buffer Insults, which uses the original vector data for accuracy.

Potential vernal pools – We used photointerpreted Potential Vernal Pools from MassWildlife's Natural Heritage and Endangered Species Program.

- Potential vernal pools that fell within a terrestrial type were treated as a single pixel pool (30 m $\times$ 30 m). When a potential vernal pool fell within a wetland mapped by DEP, we retained DEP's classification.
- Because of the inherent difficulty of identifying vernal pools from aerial photography, this layer contains many errors of commission and omission. Because there is no other data source for this important community (certified vernal pools are still quite limited and highly biased by search effort), we used these data with caveats.
- We moved vernal pools that fell in the same cell as a road over one cell to fall alongside the road in our land use. We used an algorithm that looked at the vector data to move the pool to the correct side of the road.

MassGIS networked hydro centerlines, NHD stream network – We used the MassGIS networked stream centerlines for the mainland, and filled in the Cape and islands with edited versions of NHD centerlines.

- We edited these data to repair a significant number of breaks in the network, as the CAPS watershed metrics require a connected network.
- We deleted the dense (and meaningless) network of channels in cranberry bogs, instead connecting streams flowing through bogs with straight lines passing through the bogs. These dense channels made it impossible to represent flow in a 30 m grid.
- We deleted the ditches in salt marshes for similar reasons. Streams that flowed into DEP salt marshes were retained, and any stream that originated within a salt marsh was deleted.
- We extended stream mouths all the way to the ocean
- We added stream centerlines to our landcover grid in areas that were mapped as uplands to represent smaller (1st and 2nd order) stream communities
- We burned stream centerlines into the flow grid to force streams and rivers to prevent small DEM errors from misdirecting streams and rivers
- Because stream centerlines were digitized at varying densities, resulting in bias that affected our aquatic connectedness metric, we dropped all streams with a watershed of

less than 30 ha. This has the effect of removing parts of the smaller headwater streams throughout, while making stream density more consistent.

MassGIS 5 m Digital Elevation Model – The DEM is the basis of several of our terrain-based settings variables, and of the flow used for our watershed metrics. We used the 5 m DEM because its accuracy, consistency, and overall quality was much higher than the older 30 m DEM and the DEM from the National Hydrography Dataset. We used the DEM to create a flow direction grid, the source of flow accumulation, $CaCO_3$, and watershed metrics. The DEM was also used to model tides and tidal restrictions, solar exposure, and the slope and gradient settings variables.

- We sampled this DEM up to 30 m for all analyses
- For flow modeling (flow-based settings variables and watershed metrics), we filled depressions in the DEM

Flow direction – The D8 (single-direction) flow direction grid was derived from the depressionless 30 m DEM. We then burned stream centerlines into the flow grid to ensure that stream and river beds are represented correctly. This was an iterative processes that entailed finding loops introduced by errors in stream centerlines, correcting them, and repeating the process. Flow direction is used for watershed metrics, and also for the $CaCO_3$ content settings variable.

Flow accumulation – We built a FD8 (multiple-direction) flow accumulation grid from the flow grid and DEM. This process allows a cell to flow to multiple downslope cells, giving much more realistic flow patterns in midslopes. Flow accumulation is used for the settings variables wetness and flow volume.

- We estimated the watershed area of streams flowing into Massachusetts from NHD's flow accumulation grid.

Aquatic resistance – We modified the approach of Randhir et al. 2001 (Forest Ecology and Management 143:47-56) to build a time-of-travel grid for each cell in the project area, based on land cover, slope, flow, and stream gradient. This grid was used to define the influence area within the watershed of each point for our watershed metrics.

Dams – Dams are from the Massachusetts Office of Dam Safety.

- Dams include a structural height, which we use as a surrogate for the hydrological disruption and aquatic barrier effect of each dam.
- Dam data are old with many quality issues.

Protected open space – We selected permanently protected open space (lev_prot = "P") from the MassGIS Protected and recreational open space layer. These data are used to adjust road traffic rates in parks and state forests.

Roads and road traffic – Roads and road traffic are from the MassDOT 1:5000 Roads layer (via MassGIS). Roads were reclassified into five types (expressway, primary highway, secondary highway, light duty road, and unpaved road) based on original road classes as well as surface type (for unpaved roads). For traffic rates, we used ADT, the annual average daily traffic, which MassDOT interpolated from measured traffic rates.

- Road linework and classes are generally very accurate. The primary issue is that many paper and discontinued roads are represented.
- Traffic rates are modeled from point data and thus suffer from considerable uncertainty. Traffic rates for larger roads are probably fairly accurate, as traffic on most major roads is measured on a regular basis. Unmeasured local roads are typically assigned a rate of 100 (or sometimes 200) cars/day, which is often a wild overestimate. Many of the smallest roads, including discontinued roads and new subdivision roads were assigned a traffic rate of 0. Many closed, gated, and discontinued roads were assigned traffic rates of 100. We made the following changes to traffic rates to improve their accuracy:
  - ❖ We set a minimum traffic rate of 10 cars/day for all zero-traffic roads
  - ❖ We changed the traffic rate for all unpaved and unknown type roads with traffic rates ≤ 200 that run through permanently protected open space to 10. This fixes the wild overestimates of traffic through parks and state forests such as Myles Standish, the Quabbin, and many other state forests and other large conservation areas.

Railroads – From the MassGIS trains layer. Railroads were mapped in three classes: railroad, abandoned railbed, and rail trail.

- We deleted linework where abandoned rails were shown underwater in the Quabbin, and where railroads run through major tunnels (the Hoosac tunnel)
- We integrated rail traffic into our traffic rates layer by assigning traffic rates for commuter, passenger, and freight lines, based on estimates of the number of cars for each railroad type, estimated average number of daily freight and commuter trains, and number of daily passenger trains from schedules, and expert team assignment of the relative impact of a train car to an automobile.
- We estimated the number of tracks in each rail line from GIS data for use in the terrestrial barriers settings variable.

Road-stream crossings – Bridge and culvert locations were estimated from the intersections of stream centerlines with road and railroad linework. For each crossing, we estimated both aquatic and terrestrial passability scores based on UMass Extension’s Stream Continuity Project.

- Bridges and culverts on small, unmapped streams and unmapped roads are omitted.
- Bridges and culverts on streams with < 30 ha watersheds are omitted due to stream trimming.

- Crossing scores are based on a modeling approach with rather wide confidence intervals; furthermore, models of target scores are based on expert opinion rather than empirical passability data.
- Aquatic passability is used for the aquatic barriers settings variable.
- Terrestrial passability is used in the connectedness metric to allow connectivity under roads and railroads at stream crossings.

Tidal regime – Tidal regime is estimated from a logistic regression of DEP wetlands salt marshes vs. uplands from the DEM and interpolated tide range from 120 NOAA tide stations. This gives a grid depicting the expected tidal influence at each point. Tidal regime is used as a settings variable, and also as an input to the tidal restrictions metric.

Tidal restrictions – Potential tidal restrictions are identified from the intersection of stream centerlines in the coastal area with road and railroad linework. We modeled the severity of tidal restrictions based on 75 measured tidal restrictions from field work done by CZM and DEP using a regression of the ratio of the area of expected salt marsh above each restriction (areas where the tidal regime suggest salt marshes) to the area of salt marsh mapped by DEP above each restriction. This regression was applied to all potential tidal restrictions to estimate restriction height at each point.

- Many restrictions occur where there are no roads or railroads, for instance at tide gates; we were unable to capture such restrictions.
- The relationship between the hydrological height of a restriction and the magnitude of ecological effect is unclear; however, our regression uses a measure of ecological effect (loss of salt marshes) as our target.
- The regression was significant ($P < 0.001$), but it has only moderate predictive power ($r^2 = 0.41$).

Imperviousness – Impervious surfaces are from MassGIS's impervious surface layer, based on 2005 orthophotos. This layer is at 1 m resolution. For the imperviousness settings variable, we summarized these data to percent impervious in each 30 m cell.

- Sandy or rocky areas (especially gravel pits) are often misidentified as impervious.
- Roads were included in MassGIS's impervious surface layer, so all mapped roads are identified as impervious, including paper roads, unpaved roads, and discontinued roads.

Soils – Soil depth, pH, and texture are from NRCS digital soil maps. Most of these data were mapped at 1:25,000, but only low-resolution (1:250,000) data were available for Franklin and Plymouth counties.

- Soil texture was classified on an ordinal scale of 1-6, where 1 is organic, and 2-6 range from fine to coarse textured. Values were lumped from on text classes such as "silt loam" or "very fine sandy loam." Soil texture was not supplied for open water or urban areas.

- Soil pH was based on the representative pH for each soil type. pH values are fairly coarse, with missing values for urban areas, open water, and many other areas. Soil pH for Franklin County was dropped and left missing due to poor data quality.
- Soil depth is the expected depth to bedrock, dense, or cemented layers. We log-transformed soil depth for the soildepth settings variable. Soil depth is missing for open water and is set to zero for some mountainous areas and other apparently arbitrary areas.

Calcium – We used the geology field from TNC's Ecological Land Units for calcareous and moderately calcareous near-surface bedrock. (The original source for this dataset is USGS, available on MassGIS; TNC has reclassified lithology). Our $CaCO_3$ settings variable uses these values directly for terrestrial areas, and uses a flow accumulation model for wetlands and open water.

- Lithology is mapped at scales ranging from 1:100,000 to 1:500,000, so fine details and smaller inclusions are omitted, and spatially accuracy is poor.

Wind speed, wind power – Wind speed and power data are modeled by TrueWind Solutions LLC. Wind speed is available from MassGIS; we obtained wind power (in 16 cardinal directions) from the UMass Wind Energy Center. Original data are at 200 m resolution; we downsampled these data to 30 m by interpolation.

- Our wind exposure settings variable is based on wind speed at 30 m.
- The wave exposure settings variable uses directional window power at 50 m combined with reach to estimate potential wave exposure along the coast.

Salinity – Our salinity settings variable has three classes:

- Saltwater: areas mapped in DEP wetlands as ocean, tide flat, and rocky intertidal.
- Brackish: areas mapped in DEP wetlands as salt marsh, areas mapped as open water tidal, brackish, or salt pond (poly_code = 9) in DEP wetlands, and additional areas photointerpreted as brackish for this project by Mike McHugh at MassDEP.
- Freshwater: anything that's not saltwater or brackish.

Public beaches – MassGIS's marine beaches. Used for the beach pedestrians metric.

Beach off-road vehicles – Area where off-road vehicles congregate and park on beaches were mapped by Nathalie Regis and Mike McHugh at MassDEP for this project. Areas of intensive ORV use on beaches were mapped based on information from DEP and CZM personnel as well as photointerpretation of beaches in the MassGIS DPH marine beaches layer. These data are used for the Beach ORVs metric.

Recreational beach parking lots – Parking lots for access to recreational beaches were photointerpreted by Nathalie Regis and Mike McHugh at MassDEP for this project. All parking lots that appeared to serve recreational beaches mapped in the MassGIS marine beaches layer were delineated. Data were modified based on review by experts at MassDEP and CZM. These data are used for the beach pedestrians metric.

Salt marsh ditches – Ditches in salt marshes were photointerpreted for this project by Nathalie Regis and Mike McHugh. This layer is used for the salt marsh ditching metric.

Coastal structures – Seawalls, jetties, groins, bulkheads, revetments, and breakwaters were originally obtained for most of the Massachusetts coast in field surveys by CZM. These surveys omitted some areas where access was not feasible. Omitted areas were completed based on photointerpretation by Nathalie Regis and Mike McHugh using orthophotos and oblique aerial photography. Coastal structures are used in the coastal structures metric.

Minimum winter temperature, growing season degree-days – Temperature data were obtained by downscaling modeled PRISM weather data via interpolation. Data are 30 year normals centered on 1985.

- The minimum winter temperature settings variable is the minimum of the coldest day in January or February.
- The growing season degree-days settings variable is based on the sum of monthly mean temperatures above a threshold of 10˚ C and below a threshold of 30˚ C.

Ecoregions – EPA ecoregions for Massachusetts are from MassGIS. We modified these data slightly to include all coastal cells. Ecoregions are used for IEI rescaled by ecoregions (IEI-E) and integrated IEI (IEI-I).

Watersheds – Major watersheds are from MassGIS. We modified these data slightly to include all coastal cells. Watersheds are used for IEI rescaled by watersheds (IEI-W) and integrated IEI (IEI-I).