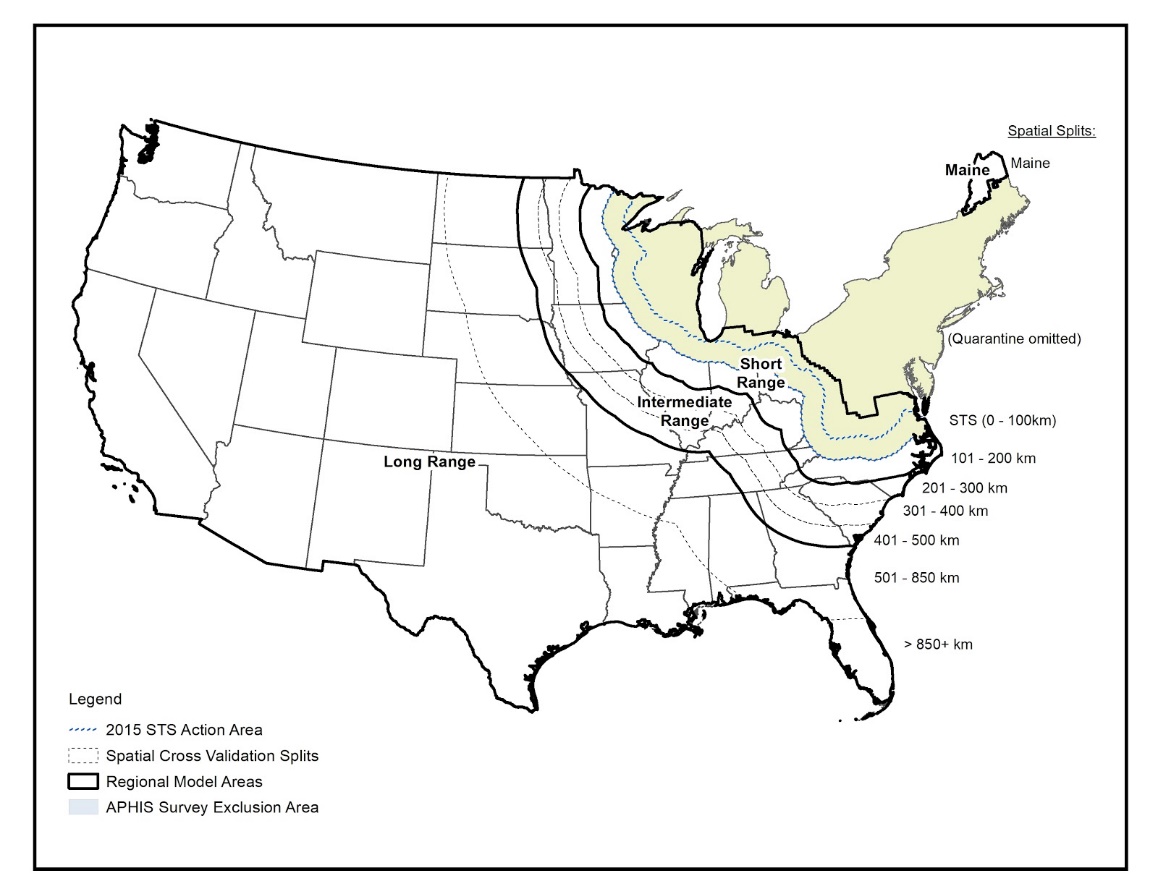
Regionalization by Pathways Importance

The Interagency Gypsy Moth working group advised us that dynamics of gypsy moth spread differed in different regions of the U.S. and that they needed a model that represented western region dynamics as well as invasion front spread dynamics. In 2013, we regionalized model development by management program area (APHIS vs Slow the Spread) to produce a detection likelihood model for the 2014 survey year [113]. The short range model was defined by data collected by the Slow the Spread program (region encompassing the active spread region and 100-km ahead of the front). The U.S. west of the short range area comprised the second regional model (i.e. the long-range model). The short range model was a pure spread kernel model, parameterized by the distance between 2011 and 2012 detections (P(detection)  = e(0.5192 - (0.00004531\*d))/(1+e(0.5192 - (0.00004531\*d)), where d=distance). The long-range model was parameterized using only anthropogenic predictors and all historical detections in this region. Due to its isolation and small geographic footprint, Maine was assumed to have median risk (sensitivity=specificity threshold) from the combined detection likelihood model.

The next model (developed in 2014 for the 2015 survey year) signaled a significant change in model regionalization approach from geographic division by management program area to geographic division by pathways predictor strength. We buffered the eastern boundary of the 2015 STS action area (i.e. the invasion front near the federal quarantine area) with 100-km spatial buffers (Figure S1), matching gypsy moth natural dispersal range. The first 0-100 km spatial buffer was equivalent to the STS action area. Buffers were combined if inadequate positive observations were available for model training, i.e. if the software returned an error for no positive observations within at least one cross-validation split. The isolated region in northern Maine was its own spatial data split during model development.



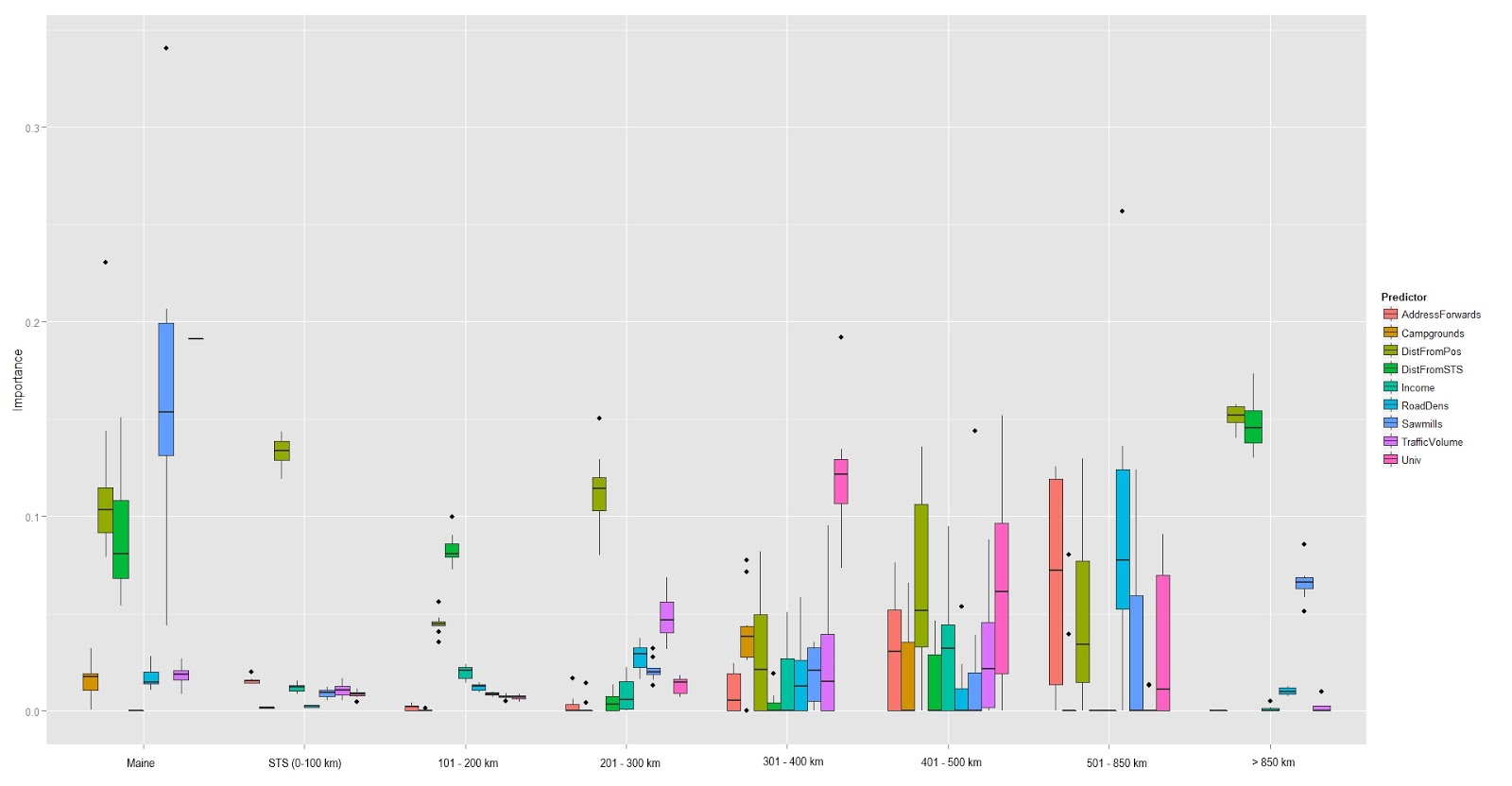
**Figure S1.** Regionalization of the study area showing the combination of individual cross-validation splits to regional model areas. The area in grey depicts the federal quarantine area where survey activity does not occur.

First, we executed a spatial stationarity (i.e. spatial cross-validation) test to evaluate whether a global model was robust to spatial variation. The 100-km buffers served as the spatial data splits for the cross-validation analysis. This process was equivalent to Miller’s [114] and Fotheringham’s [115] description of comparing local means to a global mean. Spatial cross-validation tests demonstrate that mechanisms for gypsy moth spread are not stationary across geographic space (Figure S2), with AUC values from random cross-validation being much higher than those from spatial cross-validation.  Comparison of typical random versus spatial subsetting for cross-validation demonstrates the importance for testing model robustness to spatial variation. Since a single global model was not robust to geographic variation, it was necessary to regionalize the model.

|  |  |
| --- | --- |
| 1. Random cross-validation | b)  Spatial cross-validation |
| https://lh6.googleusercontent.com/VyCaxH4yRBmueZ3uNqXLcPo2mrGyR3uD-IIHXg6cWb0MDPz2mt4x0DUCxbUALlA2gGaLFoehwvZV3KgCAnFAllWx5nyW7HkAqb2ZXGp7X4g7VPJUCNcCMUHYew-LEn27K3J_s_en | https://lh3.googleusercontent.com/adrouu0vrYCCRkXFwxkn5JC-ZtfGnqnygmJSm_B6DYD-BOSq3jn7n5iBbC_MmS-9gyPTzwOQ95kRpJmIaTU92bQpsnC-4ZMdrutAAqLi5FN8xxh88aXy0lE85y99a34BM8HNGRZq |

**Figure S2.** Results of cross-validation analyses from the 2015 model to test model robustness to various sources of variation via a) random sub-setting, or b) spatial sub-setting.

Once we established that a global model was statistically inappropriate, we proceeded with model regionalization using the spatial splits. We ran individual models for each spatial buffer using the same set of predictors, and a random cross-validation to evaluate variability in predictor response. We used a pairwise Wilcoxon rank sum test with a Holm-Bonferonni multiple comparison correction [116] in R [117] to evaluate significant shifts in pathways’ predictor strength. Pairwise significant differences were only evaluated in contiguous spatial splits, not simultaneously across all spatial splits. The selection of geographic division(s) for regionalizing the gypsy moth risk assessment was based on the consistency of significant geographic breaks across all variables, weighted by their relative importance. Analysis of variable predictor importance over space shows that all predictor variables expressed significant non-stationarity, and there was considerable variation in the geographic breaks across variables (Figure S2). There was also increasing uncertainty in the importance of the predictor variables as the model was trained on different regions based on spatial splits increasingly removed from the generally infested area. This is due to the decreasing pest prevalence and the random assignment of positive detections to cross-validation splits. There was also a shift in importance shifted from the biological dispersal predictors to the anthropogenic predictors around 200 km.  However, biological and anthropogenic introduction variables continued to interact up to 500 km away from the spread front.



**Figure S3.** Box plots of variable importance for the models for individual spatial splits varied across geographic space for the 2015 model, showing variation between the 10-fold cross-validation results for each spatial split and uncertainty increased in spatial splits increasingly far from the infested area.

Analysis of the Wilcoxon rank sum tests for significant grouping, as well as inspection of individual variable importance plots across space (Figure S2), revealed some interesting trends. Beginning with the biological dispersal predictors, the distance from the STS zone expressed an expected behavior of declining importance across space, with the exception of the final spatial buffer (>850 km). The other biological dispersal parameter, distance from prior year positive, showed considerable variation in importance over space. This is likely due to the predictor representing localized spread, and likely uninformative to regionalizing the model. The importance of address forwarding data (proxy for human-assisted, long distance dispersal) increased with distance from the spread front. The Maine data split became a stand-alone region, the STS project area (0-100km) and 101-200 km data splits were grouped to a “Short Range” geographic region [60], the 201-500 km data splits were grouped to an “Intermediate Range” geographic region, and the remaining data splits were grouped into a “Long Range” geographic region.  Due to the historical nature of the STS action area, some data is available into the current federal quarantine area. Therefore, we included historical regions of the STS project area (beginning in 2001). We ended with four distinct model regions: a Maine-only model because it was geographically isolated, a short-range model, an intermediate-range model, and a long-range model. This stationarity analysis was re-executed in subsequent model years with no significant change in results. Therefore, the model regions identified in the 2015 model remained constant for the 2016 and 2018 model years (Figure S1).