5-2014

Application of Spatial Modeling Tools to Predict Native Bee Abundance in Maine's Lowbush Blueberries

Shannon J. Chapin

THE APPLICATION OF SPATIAL MODELING TOOLS TO PREDICT NATIVE BEE ABUNDANCE IN MAINE'S LOWBUSH BLUEBERRIES

By

Shannon J. Chapin

B.S. The Pennsylvania State University, 2007

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Ecology and Environmental Science)

The Graduate School

The University of Maine

May 2014

Advisory Committee:

Cynthia S. Loftin, Unit Leader, U.S. Geological Survey Maine Cooperative Fish and Wildlife Research Unit, Associate Professor of Wildlife Ecology, Co-Advisor

Francis A. Drummond, Professor of Insect Ecology and Insect Pest Management, Co-Advisor

Brian McGill, Associate Professor of Biological Science

THE APPLICATION OF SPATIAL MODELING TOOLS TO PREDICT NATIVE BEE ABUNDANCE IN MAINE'S LOWBUSH BLUEBERRIES

By Shannon J. Chapin

Thesis Co-Advisors: Dr. Cynthia S. Loftin

Dr. Francis A. Drummond

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Ecology and Environmental Science)
May 2014

Non-native honeybees historically have been used to pollinate many crops throughout the United States, however, recent population declines have revealed the need for a more sustainable pollination plan. Native bees are a natural resource that can play an important role in pollination. I used spatial modeling tools to evaluate relationships between landscape factors and native bee abundance, with a focus on the wild native bees that pollinate Maine's lowbush blueberries. I applied the InVEST Crop Pollination ecosystem spatial modeling tool, which predicts pollinator abundance based on available floral resources and nesting habitat, to the Downeast Maine region. The InVEST model is a generic tool that can be adapted to any landscape with development of location specific parameters and a validation dataset. I surveyed botanists, entomologists and ecologists who are experts in native bee ecology and familiar with Maine's landscape, and asked them to rank the suitability of landcover types as native bee habitat. I used previously

collected bee abundance data to validate model assumptions. I evaluated the sensitivity and explanatory power of the InVEST model with four model parameterization methods: 1) suitability values assigned through the expert survey; 2) suitability values developed through a sensitivity analysis; 3) informed suitability values developed through an optimization based on the sensitivity analysis; and, 4) uninformed suitability values developed through machine-learning simulated annealing optimization. I evaluated the improvement in prediction gained from expert-informed and optimization-informed parameterization compared with prediction based on the relationship between proportion of landcover surrounding blueberry fields and native bee abundance as an alternative to the InVEST model. The InVEST model parameterized through expert opinion predicted native been abundance ($r = 0.315$; $P = 0.047$), whereas, the uninformed optimization improved model performance by 28% ($r = 0.404$; $P = 0.010$), and the informed optimization technique improved model performance by 58% ($r = 0.486$; $P = 0.002$). The landcover analysis found a significant relationship between the proportion of deciduous/mixed forest within a 2000 meter buffer around a field and native bee abundance within the field ($r = 0.446$; $P = 0.004$). Although the InVEST model reliably predicts bee abundance across a landscape, simpler models quantifying relationships between bee abundance and proportional land cover around focal fields may be suitable alternatives to the InVEST simulation model.

ACKNOWLEDGMENTS

Financial support for this project was provided by the U.S. Department of Agriculture National Institute of Food and Agriculture - Specialty Crops Research Initiative Contract/Grant/Agreement No. 2011-51181-30673, U.S. Geological Survey Maine Cooperative Fish and Wildlife Research, and the University of Maine Department of Wildlife Ecology.

Many thanks go to my advisors, Cynthia Loftin and Frank Drummond, and my committee member Brian McGill, for providing guidance and support.

I would also like to thank Sara Bushmann for allowing me to use her data to validate model predictions. Additional thanks go those who participated in the expert opinion survey (approved through the Institutional Review Board - Application # 2013-0506), and many additional folks who provided feedback related to my project.

Thank you also to the Pollination research team, the Loftin lab, students and faculty in the Wildlife Ecology department and Ecology and Environmental Science program and family and friends.

Finally, I would like to thank my husband, Luke, for being a great cheerleader, and distraction, when needed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Nearly 75% of the world's crops rely at least partly on animal pollination (Klein et al. 2007), and bees are the most important insect pollinator (Tepedino 1979). Lowbush blueberry (*Vaccinium angustifolium*), a leading crop industry in Maine, requires insect pollination (Drummond 2002). Maine is the world's second largest producer of wild blueberries with over 91.1 million pounds harvested in 2012 (Yarborough 2012) and the country's second largest importer of non-native honeybees (*Apis mellifera*) for pollination, with more than 75,000 hives deployed yearly (A. Jadczak, Maine Department of Agriculture, pers. comm.). Maine manages the greatest area (>24,000 ha) in lowbush blueberries of any state (Yarborough 2009), primarily in Hancock and Washington counties. The decline of honeybee populations has increased the cost of hive rentals (Pettis and Delaplane 2010). Focus increasingly has turned to a more sustainable pollination plan, which includes partially relying on and improving populations of native pollinators. Native pollinators provide a freely available ecosystem service. They have coevolved with wild lowbush blueberries, and they are adapted to forage in reduced light and cooler temperatures common where blueberries grow (Cane and Payne 1988).

There are more than 270 native bee species in six families *(Andrenidae, Apidae, Colletidae, Halictidae, Megachilidae, Melittidae)* in Maine (Drummond and Stubbs 2003, Dibble et al. unpublished data). More than 40 bee species forage in lowbush blueberries in Maine, although there likely are more associated species, as > 60 species have been recognized on blueberries in Nova Scotia (Drummond and Stubbs 2003). While these families exhibit various life history traits, all require at a minimum, two key components to survive: suitable nesting habitat and floral resources for forage (Lonsdorf et al. 2009, 2011).

The proportion of natural habitat surrounding a crop field affects pollination by bees, as "natural habitat" can be synonymous with resources that provide nesting and foraging habitats. Specifically, in a synthesis of 29 studies examining pollination services, Garibaldi et al. (2011) determined that bee visitation to crop bloom decreased with isolation from natural areas, despite added honeybee visits. The definition of natural habitat varies by geographic location and at its simplest includes environments that offer shelter, nesting grounds and food resources (Ricketts et al. 2008). Natural habitat that provides nesting and foraging resources to bees in Maine is represented in land cover maps as deciduous/mixed forest, deciduous/mixed forest edge, and old fields and grasslands. The Downeast region of Maine, where 85% of the world's lowbush blueberries are harvested (Henly 2012), has few people (averaging 9.1 persons per square kilometer (km), compared to the US average of 33.7 (U.S. Census Bureau 2014); the predominantly rural land development includes home gardens, which may provide additional beneficial habitat for bees.

The InVEST Crop Pollination Model, developed by the Natural Capital Project (Lonsdorf et al. 2009, 2011), is a tool for examining relationships between relative bee abundance and landscape composition. Bees are mobile organisms that depend on resources that often vary spatially and temporally across a landscape (Kremen et al. 2007), and access to those resources depends on the foraging ability of the bee (Patricio-Roberto and Campos 2014). Understanding factors affecting pollination services on a farm requires understanding relationships between the spatial distribution of pollinator habitat surrounding a farm and bee abundance in the focal crop (Kremen et al. 2007). The InVEST Model predicts relative abundance of pollinators across a landscape, based on nesting resources within the focal cell and floral resources surrounding the cell within the confines of the modeled bee's foraging range. InVEST can be adapted to any crop, however, it requires validation for the focal crop.

The InVEST model requires a spatial landcover dataset and parameters relating landcover suitability for providing habitat resources given the modeled bee's life history strategy (Lonsdorf et al. 2009, 2011). In the absence of empirical data, parameters can be assigned based on published values or expert opinion. Expert opinion often is used to inform spatial models (Compton et al. 2007, Lonsdorf et al. 2009, Spear et al. 2010, Kennedy et al. 2013), although predictive accuracy of the model is not necessarily improved with this knowledge (Charney 2012). The abundance of pollinators may be affected not only by landscape composition, but also by the pattern and arrangement of the surrounding landscape (Brosi et al. 2008, Ricketts et al. 2008, Lonsdorf et al. 2009, 2011) and the scale and extent at which the landscape is modeled (Lonsdorf et al. 2009, 2011).

I investigated relationships between landscape composition and native bee abundance with the InVEST Crop Pollination Model adapted to Downeast Maine's landscape, with lowbush blueberry fields as the focal study system. My analyses addressed the following questions: 1) Does expert opinion ranking of bee habitats (the most common parameterization technique used for InVEST) provide predictive capability for estimating bee abundance? 2) How does the predictive capability of the InVEST model compare across several parameterization techniques? 3) How do predictions of a simple proportional landcover model compare to those of the InVEST model?

DESCRIPTION OF STUDY EXTENT AND FIELD SITES

I applied the InVEST model to the 4,802 km^2 blueberry growing region containing 40 focal blueberry fields (< 1 - 17 ha) in the Downeast region of Maine (Figure 1.) Additional applications of the InVEST model across different extents are described in Appendix B.

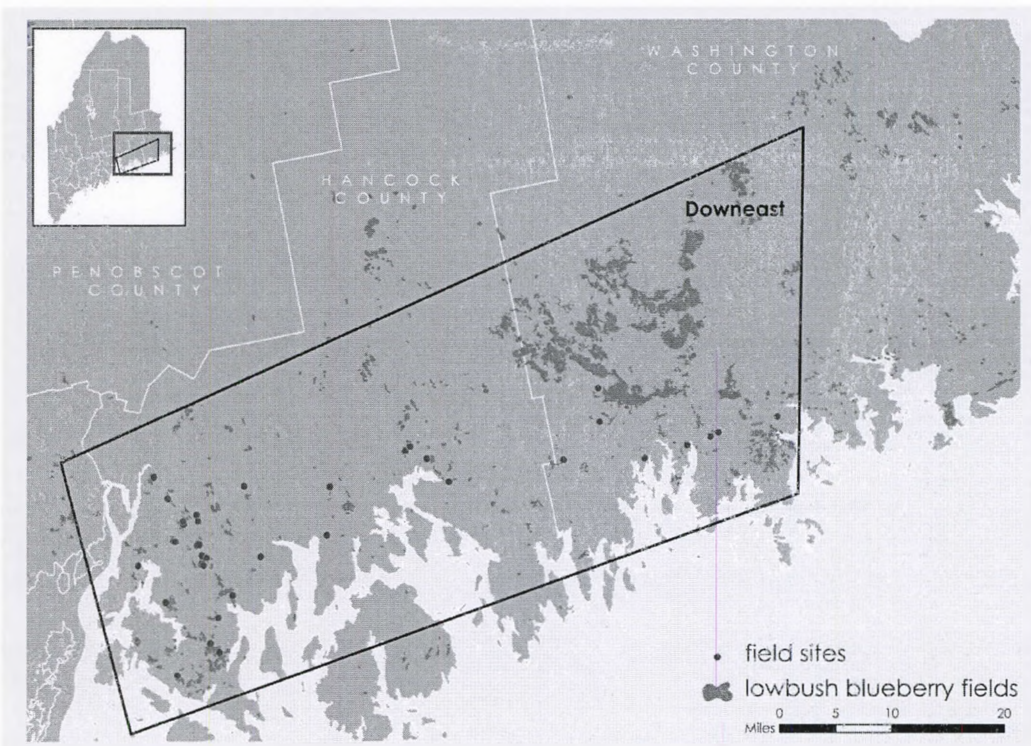

Figure 1. Modeled extent and blueberry field sites used for validation of the InVEST model in the Downeast region of Maine, USA.

METHODS

Spatial landcover dataset selection and processing

The InVEST model requires an accurate spatial landcover dataset. The Maine Landcover Dataset 2004 (MELCD 2004; http://www.maine.gov/megis/catalog/)

combines the National Landcover Dataset 2001 (NLCD 2001), based on 1999-2001 Landsat Thematic Mapper 5 and 7 imagery, with classification of 2004 SPOT 5 imagery, to create a 5-meter resolution raster dataset with 23 landcover classes. The blueberry field category represents commercial blueberry operations with an accuracy of 89.7% in Maine.

I updated the 2004 MELCD landcover layer with ancillary datasets (ArcGIS ® version 10.0; Environmental Systems Research Institute, Redlands, CA, United States), including railroads (RAILROUTESYS) and roads (MEDOTPUBRDS, NG911; http://www.maine.gov/megis/catalog/). I updated the MELCD wetlands classes (wetland forest, wetlands, scrub-shrub) with the National Wetland Inventory (NWI; http://www.maine.gov/megis/catalog/) to capture wetland diversity potentially important to foraging bees. I created a *deciduous/mixed forest edge* class by applying a 10m buffer around *deciduous forest* and *mixed forest* pixels. I resampled the USDA Croplands Dataset (CDL; http://nassgeodata.gmu.edu/CropScape/) to 5-m pixels, and I updated the MELCD "blueberry field" class with blueberry fields >4 hectares in the CDL, capturing fields omitted from the original MELCD dataset while excluding wild blueberries not in managed fields. I digitized the perimeter of blueberry fields where bee samples were collected but that were missing from the compiled landcover dataset. The final landcover 5-m pixel dataset reclassified 42 classes into eight landcover types: *deciduous/mixed forest edge*, *developed/other*, *coniferous forest*, *deciduous/mixed forest*, *emergent/shrub-shrub wetlands*, *other wetlands/water*, *agriculture/field* and *blueberries*.

Bee species life history parameterization

I modeled 14 solitary bee species (Table 1.) in four families representative of the lowbush blueberry solitary bee community (Bushmann 2013). I assigned life history parameters (i.e., nesting preferences, flight seasonality) based on expert opinion and literature references (Osgood 1972, Michener 1966, Cane 1992, Michener 2000, Asher and Pickering 2013; Table 1.).

Table 1. Life history traits of modeled solitary bee species.

Species	Family	Nest substrate	Typical foraging distance (m)	Flight season
Andrena carlini	*Andrenidae*	ground	598	Mar - Aug
Andrena carolina	*Andrenidae*	ground	246	Apr - Jul
Andrena vicina	*Andrenidae*	ground	569	Mar - Aug
Augochlorella aurata	*Halictidae*	ground	60	Apr - Oct
Colletes inaequalis	*Colletidae*	ground	1091	Mar - Sept
Halictus ligatus	*Halictidae*	ground	148	Mar – Nov
Lasioglossum acuminatum	*Halictidae*	ground	186	Apr – Oct
Lasioglossum cressonii	*Halictidae*	cavity	63	Mar – Oct
Lasioglossum heterognathum	*Halictidae*	ground	16	Apr - Sept
Lasioglossum leucocomum	*Halictidae*	ground	31	Mar - Oct
Lasioglossum pectorale	*Halictidae*	ground	81	Mar – Nov
Lasioglossum versatum	*Halictidae*	ground	79	Mar – Oct
Osmia atriventris	*Megachilidae*	cavity	186	Apr – Jul
Osmia inspergens	*Megachilidae*	cavity	495	May – June

Foraging estimates obtained from inter-tegular width measurements

I estimated foraging ranges of locally captured bees by measuring the inter-tegular (IT) width (i.e., distance between the wing bases) with a Dino-Lite mobile digital microscope and analyzed images in Dino-Capture 2.0 (AnMo Electronics Corporation, Hsinchu, Taiwan). I estimated foraging ranges from the measured IT width (mm) based on regression formulae developed by Greenleaf et al. (2007). Five measurements were taken per specimen, and 10 specimens were measured per species, except for *Osmia atriventris*, with only eight specimens available (Figure 2.). I averaged the measured IT widths by species (n = 50; n = 40 for *O. atriventris*), and I calculated both maximum and typical homing distances (m) (Table 2.) (Greenleaf et al. 2007). Mean typical homing distance values per species were used for model parameterization (Table 1.).

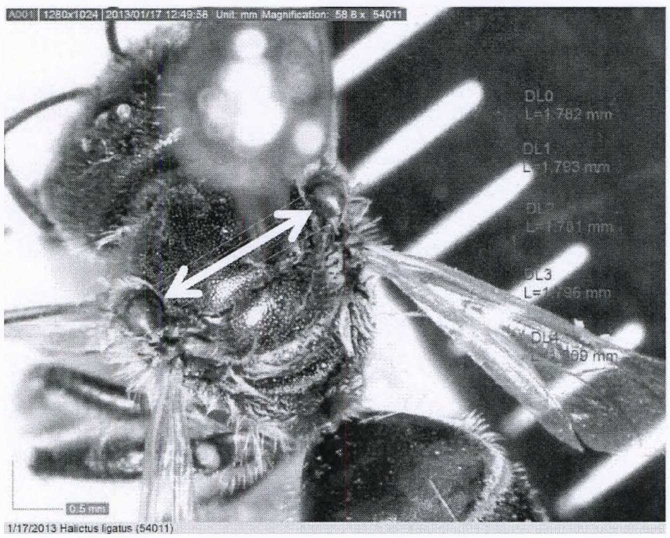

Figure 2. Example of the IT measurements used to estimate foraging distance (*Halictus ligatus*).

Table 2. Mean (±standard deviation) measured IT widths and mean typical and maximum homing distances.

Species	Mean IT width (mm)	Mean typical homing distance (m)	Mean maximum homing distance (m)
Andrena carlini	2.74(0.14)	598	1290
Andrena carolina	2.08(0.09)	246	513
Andrena vicina	2.70(0.17)	569	1226
Augochlorella aurata	1.35(0.10)	60	120
Colletes inaequalis	3.30(0.19)	1091	2410
Halictus ligatus	1.78(0.16)	148	302
Lasioglossum acuminatum	1.91(0.09)	186	385
Lasioglossum cressonii	1.37(0.12)	63	125
Lasioglossum heterognathum	0.91(0.18)	16	31
Lasioglossum leucocomum	1.10(0.10)	31	59
Lasioglossum pectorale	1.48(0.13)	81	162
Lasioglossum versatum	1.47(0.15)	79	157
Osmia atriventris	1.91(0.20)	186	384
Osmia inspergens	2.59(0.17)	495	1060

Landcover suitability parameterization through expert survey

I derived estimates of the suitability of landcover types for both floral and nesting habitat for bees from an expert survey of 16 entomologists, ecologists, and botanists familiar with landscapes in Maine. The experts ranked (0=unsuitable to 10=most suitable) landcover class suitability for ground and cavity nesting bees, and spring, early summer and late summer forage (Appendix A). Participants responded either to a printed survey distributed by the US Postal Service or an electronic survey distributed by email. I summarized survey responses by class range, mode and average, omitting the *coniferous forest - clearcut* landcover type in my models as I did not have access to a current spatial landcover that represented that type. I rescaled responses (1-10), and used the average scaled response as the suitability ranking for the landcover or nesting substrate. I divided the average scaled suitability values by 10 to meet the InVEST model parameter range requirement of 0.1 - 1.0 (Table 3.).

Table 3. Average (± standard deviation) scaled landcover suitability values assigned through expert opinion.

Landcover	Ground nesting	Cavity nesting	Spring forage	Early Summer forage	Late Summer forage
Deciduous/mixed forest, edge	0.9(0.17)	1.0(0.19)	0.9(0.24)	0.9(0.24)	1.0(0.22)
Developed/other	0.9(0.25)	0.6(0.30)	1.0(0.27)	0.9(0.26)	1.0(0.22)
Coniferous forest	0.5(0.23)	0.6(0.28)	0.1(0.24)	0.1(0.21)	0.1(0.29)
Deciduous forest/mixed forest	0.6(0.21)	0.9(0.22)	0.7(0.21)	0.5(0.29)	0.4(0.18)
Emergent wetlands/scrub-shrub	0.2(0.14)	0.4(0.24)	0.7(0.22)	0.6(0.25)	0.6(0.20)
Wetlands/water	0.1(0)	0.1(0.05)	0.3(0.20)	0.2(0.16)	0.5(0.18)
Agriculture/field	0.7(0.29)	0.2(0.18)	0.9(0.31)	0.7(0.27)	0.9(0.33)
Blueberries	1.0(0.25)	0.4(0.26)	0.4(0.29)	1.0(0.28)	0.5(0.26)

InVEST model parameterization

I evaluated the sensitivity and explanatory power of the InVEST model with four model parameterization methods: 1) suitability values assigned through expert opinion, 2) suitability values developed through sensitivity analyses, 3) suitability values developed through informed optimization, and 4) suitability values developed through uninformed simulated annealing optimization. The InVEST model was applied to 14 focal species, and all models were validated with bee data collected from 40 fields during 2010-2012 (Bushmann 2013). Though the resolution of the final updated dataset remained at 5-m, I conducted the InVEST analysis at a 10-m resolution to decrease analysis time.

Expert opinion

I ran the InVEST model with average suitability values resulting from the expert survey. I evaluated the relationship between the InVEST model output and the field-collected bee abundance data with simple linear regression and Pearson product moment correlation coefficients (R 2.14.1, R Development Core Team 2011). I compared the three parameterization techniques to results from this baseline model.

Sensitivity analysis

I evaluated how uncertainty in parameter choice influenced the output of the model with a sensitivity analysis. I iteratively ran the model, varying each of the 40 nesting and floral resource suitability parameters individually by ± 0.1 (i.e., ± 10%) ranging 0-1; for a total of 74 model runs. Some parameters initially were assigned the maximum value =1 and therefore were not evaluated at smaller values (Table 3.).

I evaluated the relationship between the InVEST model output and the field-collected bee abundance data with simple linear regression and percent change in the Pearson product moment correlation coefficients (r) compared to the baseline model parameterized by expert opinion.

Informed optimization

I conducted an optimization of the InVEST model informed by the sensitivity analysis. I varied the number of parameters altered and the amount of change in suitability values in nine model runs. For example, one run included 20% (0.2) decreased suitability of *blueberries* for nesting and forage, whereas, another run altered all parameters by ±20% (0.2), with direction determined by the sensitivity analysis. I evaluated the relationship between the InVEST model output and the field-collected bee abundance data with simple linear regression and the Pearson product moment correlation coefficients.

Uninformed optimization

I used simulating annealing optimization to parameterize the model with uninformed suitability values optimized to the validation dataset (Kirkpatrick et al. 1983). Simulated annealing is an optimization process that enables a function to escape local minimums and local maximums, with the goal to instead find a global optimum. The function is able to move both uphill and downhill, first with large jumps, and then with subsequent smaller jumps as the function focuses in on the optimum (Goffe et al. 1994).

This technique was performed by embedding the InVEST model into a function and running it through Python's minimizing *scip.optimize.anneal* function (Oliphant

2007). Intial input parameters were those assigned through the expert opinion survey. All parameters varied simultaneously for each run. *Scip.optimize.anneal* is a minimizing function (i.e., seeks the minimum optimal value) therefore, I set the function to attempt to maximize the correlation coefficient by multiplying it by -1 to convert the value to positive. The optimization completed 87 iterations, although it failed to identify a global minimum given computer resource limitations. I evaluated the relationship between the InVEST model outputs for each optimized run against the field-collected bee abundance data with simple linear regression and calculated the Pearson product moment correlation coefficients.

Simple proportional landscape analysis

I calculated the average proportion of landcover types in 500, 1000, 1500 and 2000 m buffers surrounding the 40 fields where bees were collected (Table 4.) to compare with bee abundance in these fields (ArcGIS v. 10.0, Environmental Systems Research Institute, Redlands, CA, United States; Geospatial Modelling Environment GME; Beyer 2012).

Table 4. Average (±standard deviation) proportions of landcover cover classes within a 500, 1000, 1500 and 2000 m buffer surrounding field sites (n = 40).

	500 m	1000 m	1500 m	2000 m
Deciduous/mixed forest, edge	0.06 (0.02)	0.06 (0.01)	0.05 (0.02)	0.05 (0.01)
Developed/other	0.04 (0.03)	0.04 (0.02)	0.04 (0.02)	0.04 (0.02)
Coniferous forest	0.29 (0.18)	0.34 (0.15)	0.35 (0.14)	0.36 (0.14)
Deciduous forest/mixed forest	0.30 (0.18)	0.28 (0.16)	0.27 (0.14)	0.26 (0.12)
Emergent wetlands/scrub-shrub	0.08 (0.07)	0.09 (0.07)	0.09 (0.05)	0.10 (0.04)
Wetlands/water	0.04 (0.08)	0.06 (0.09)	0.08 (0.10)	0.10 (0.10)
Agriculture/field	0.05 (0.04)	0.04 (0.03)	0.04 (0.03)	0.04 (0.02)
Blueberries	0.14 (0.13)	0.10 (0.09)	0.08 (0.07	0.06 (0.06)

The proportions were calculated using ArcGIS version 10.0 (Environmental Systems Research Institute, Redlands, CA, United States), and Geospatial Modelling Environment (GME; Beyer 2012). First, I used the ArcGIS "Buffer (Analysis)" tool to buffer all fields by the four selected buffer distances, and then used the GME "Intersect Polygons with Raster" tool to summarize the proportions of landcover classes within the buffer polygons.

I compared the landcover proportion in each buffer for each of the 14 species included in the InVEST model evaluation as well as for bee abundance data collected from the same 40 field sites for another 6-45 species not used in the model analysis. I evaluated the relationship between the proportion of landcover types and observed bee abundance within each field with simple linear regression and the Pearson product - moment correlation coefficient (*r*).

RESULTS

Expert survey

Twelve of 16 experts completed the survey, with 92% preferring the electronic version. Responses varied with the greatest agreement in the value of *wetlands/water*, and the least agreement in the value of *agriculture/field* (Table 5).

Table 5. Range (maximum - minimum) of expert survey derived suitability values.

Landcover	Cavity nesting	Ground nesting	Spring forage	Early summer forage	Late summer forage
Deciduous/mixed forest, edge	4	5	6	7	8
Developed/other	9	8	9	8	6
Coniferous forest	9	7	3	1	1
Deciduous forest/mixed forest	6	6	7	9	6
Emergent wetlands/scrub shrub	9	4	8	9	6
Wetlands/water	1	0	5	5	5
Agriculture/field	6	9	9	9	9
Blueberries	7	8	9	7	6

Evaluation of alternative models

Baseline InVEST model - parameterized through expert opinion

The InVEST predictions of bee abundance in the modeled Downeast extent were significantly correlated with field-collected abundances (Pearson's $r = 0.315$; $P = 0.047$). I compared the parameterization analyses to this model.

Sensitivity analysis

Altering the model parameters by ±10% resulted in a change in correlation coefficient values of -7.09 - +9.09% (Table 6.). The model is most sensitive to changes in the *deciduous/mixed forest* and *blueberries* landcover classes. Decreasing the value of all suitability parameters for the *blueberry* class resulted in increased correlations (Table 6.). An increase in the value of the ground nesting parameter, and early summer and summer floral suitability for *deciduous/mixed forest* resulted in an increase in correlation strength (Table 6.).

Table 6. Results of sensitivity analysis. Values shown are percent change in Pearson correlation coefficient (*r*) for ± 10% change in parameter value compared to the baseline model.

	Deciduous/ mixed forest, edge	*Developed/ other*	*Coniferous forest*	*Deciduous/mixed forest*	*Emergent wetlands/ scrub shrub*	*Wetlands/ water*	*Agriculture/ field*	*Blueberries*
cavity (-)	-0.46*	-0.92*	0.50*	-1.98*	0.02*	-0.38*	0.11*	3.34*
cavity (+)	-	0.90*	-0.51*	1.91*	-0.04*	-0.12*	-0.12*	-3.21
ground (-)	-1.87*	-3.42	1.47*	-7.09	-0.12*	-1.40*	0.60*	9.09*
ground (+)	1.80*	3.32*	-1.64*	6.30*	-0.04*	1.36*	-0.74*	-
spring (-)	-0.67*	-1.15*	1.24*	-2.98*	0.03*	-1.01*	0.38*	4.83*
spring (+)	0.66*	-	-1.34*	2.83*	-0.04*	1.00*	-0.40*	-4.69
early sum. (-)	-0.93*	-1.52*	1.52*	-5.03	-0.06*	-0.59*	0.62*	6.72*
early sum. (+)	0.90*	1.50*	-1.72*	4.62*	0.02*	0.58*	-0.66*	-
summer (-)	-1.08*	-1.81*	1.74*	-4.09	0.05*	-0.70*	0.42*	6.71*
summer (+)	-	-	-1.91*	3.83*	-0.09*	0.69*	-0.47*	-6.52

*model run significant at <0.05

Informed optimization

All 9 runs parameterized through the informed optimization process performed better (2.671% - 54.024%) than the expert-informed baseline run (i.e., Pearson's $r > 0.316$; $P < 0.047$). The best performing model used the majority of the expert derived parameters altered in ±0.2 in the direction the sensitivity analysis implied increased model fitness (Table 7.). This run performed 54% better than the baseline model run ($r = 0.486$; $P = 0.002$).

Table 7. Parameters used in the best performing model through informed optimization. Expert assigned parameters in parenthesis.

Landcover	Ground nesting	Cavity nesting	Spring forage	Early summer forage	Late summer forage
Deciduous/mixed forest, edge	1.0(0.9)	1.0(1.0)	1.0(0.9)	1.0(0.9)	1.0(1.0)
Developed/other	1.0(0.9)	0.8(0.6)	1.0(1.0)	1.0(0.9)	1.0(1.0)
Coniferous forest	0.3(0.5)	0.4(0.6)	0.0(0.1)	0.0(0.1)	0.0(0.1)
Deciduous forest/mixed forest	0.8(0.6)	1.0(0.9)	0.9(0.7)	0.7(0.5)	0.6(0.4)
Emergent wetlands/scrub-shrub	0.2(0.2)	0.2(0.4)	0.5(0.7)	0.8(0.6)	0.4(0.6)
Wetlands/water	0.3(0.1)	0.3(0.1)	0.5(0.3)	0.4(0.2)	0.7(0.5)
Agriculture/field	0.5(0.7)	0.0(0.2)	0.7(0.9)	0.5(0.7)	0.7(0.9)
Blueberries	0.8(1.0)	0.2(0.4)	0.2(0.4)	0.8(1.0)	0.3(0.5)

Uninformed optimization

The simulated annealing optimization of parameter values resulted in correlation coefficients ranging $r = -0.460$ to $r = 0.404$. The best performing model ($r = 0.404$; $P = 0.010$) performed 29% better than the baseline or expert-informed model.

Simple proportional landscape analysis

I observed significant positive correlations between the proportion of *deciduous/mixed forest* and bee abundance of the 14 selected species at the 500, 1500 and 2000 m buffers (Table 8.). The strongest correlation occurred with the proportion of the *developed/other* landcover class surrounding the field at both the 1500 and 2000 m scale (Table 8.). Results of other landcover classes varied in significance and strength across all scales, but the majority were constant in direction (Table 8.).

For the total dataset (sum of all taxa abundance), I observed significant positive correlations between the proportion of *deciduous/mixed forest* and bee abundance, and significant negative correlations between the proportion of *coniferous forest* and bee abundance (Table 8.). Both relationships were strongest at the 2000 meter scale.

Table 8. Pearson product-moment correlation values (r) for both proportional landscape analyses at the 500, 1000, 1500 and 2000 meter (m) scale. Correlations between landcover and observed bee abundance for 14 selected species shown first; correlations between landcover and total observed bee abundance (sum of all taxa abundance) shown second.

Landcover	500 m	1000 m	1500 m	2000 m
Deciduous/mixed forest, edge	0.20; 0.25	0.27; 0.26	0.32; 0.33*	0.34**; 0.41**
Developed/other	0.09; 0.09	0.15; 0.09	0.40**; 0.35*	0.46**; 0.40**
Coniferous forest	-0.23; - 0.32*	-0.24; -0.36*	-0.30*; -0.42**	-0.34*; -0.47**
Deciduous/mixed forest	0.31*; 0.40**	0.27; 0.36*	0.32*; 0.42**	0.34*; 0.45**
Emergent wetlands/scrub-shrub	-0.01; 0.00	0.06; 0.07	-0.12; -0.12	-0.18; -0.19
Wetlands/water	0.19; 0.28	0.15; 0.27	0.18; -0.28	0.19; 0.29
Agriculture/field	-0.29; -0.32*	-0.30*; -0.36*	-0.31*; -0.37**	-0.27; -0.31
Blueberries	-0.18; -0.26	-0.22; -0.28	-0.27; -0.32*	-0.22; -0.26

*significant at <0.05; **significant at <0.01

DISCUSSION

The InVEST model, like other spatial models that result in predictive maps, can be a powerful tool that is relatively easy to adapt to new areas. That being said, my research demonstrates that it is important to assess the effect of parameterization techniques on the predictive ability of the model.

Reliability of model predictions can be affected by the model parameterization approach; responses of predictions to changes in parameter values may reveal unexpected model behavior and outcomes. The InVEST model parameterized through informed optimization performed better than the expert-opinion informed model. This improvement was not unexpected; the optimization process is data driven, and therefore it maximizes model prediction performance by altering the parameters to best fit the data. Model parameterization with the uninformed, machine learning, simulated annealing also was more reliable than the model driven by the expert opinion survey results; this process determines the global optimum for nearly all functions (Clarke et al. 2009), with improved prediction accuracy over model performance affected by lack of agreement in parameter values revealed in the expert surveys. A simple, proportional landscape analysis had greater predictive power than the InVEST model, emphasizing that the goal and scale of the prediction are important considerations when selecting the parameterization approach.

Limitations of expert opinion

Expert opinion surveys often are used to parameterize models developed to facilitate conservation efforts (Compton et al. 2007, Lonsdorf et al. 2009; Spear et al. 2010; Kennedy et al. 2013) in two approaches: responses are first recorded independently and then combined, or the group works together to arrive at a consensus (Martin et al. 2012). My expert opinion survey had limited consensus, reflecting expert group uncertainty, of landcover suitability for nesting and foraging habitat. Given that I solicited the experts' opinions individually, there was no opportunity to reduce this uncertainty or disagreement through discussion. Between-expert uncertainty rarely is explored (Johnson et al. 2004) but can be an important contribution to model prediction error. Elicitation of independent expert parameter valuation provides an opportunity to examine effects of parameter uncertainty that can reduce bias in decision-making (Czembor et al. 2011). I parameterized the InVEST model with the re-scaled average response value (Martin et al. 2012), which relativized and generalized the values and as a result may have increased error in parameter values. The lack of empirical data of landcover suitability as native bee habitat in Maine increases reliance on expert evaluation of parameters. An expert may not accurately extrapolate their within-region knowledge to outside their area of experience; there is no opportunity to control for this error, resulting in poorly constructed predictive models (Murray et al. 2009). I selected experts familiar with Maine's landscape and native bees, although their experience was not necessarily in the area included in the modeling extent. In addition to varied expert experience, variation in the responses could reflect true variation in the landscape as many landcover classes used in the model have naturally patchy distributions of both floral and nesting resources (Cane 2001). This fine-scale diversity in the model predictions may be obscured by the model resolution. Model resolution was implicated in a previous application of the InVEST model. Specifically, the model was unable to accurately predict abundance in New Jersey; the authors speculated that this was due to the coarse (i.e., 30 m) landcover layer used not capturing fine scale heterogeneity present in the landscape (Lonsdorf et al. 2009, 2011).

Studies to quantify suitability and bee use of the variety of habitats in Maine will be improved with robust parameterization based on empirical data. Additionally, the potential for an expert panel to provide values reached through consensus would be beneficial to explore (Kennedy et al 2013).

Sensitivity across parameters

The InVEST model was most sensitive to changes in the suitability ranking of *deciduous/mixed forest* and *blueberries* landcover classes. *Deciduous/mixed forest* is a dominant land cover type surrounding blueberry fields, and model sensitivity to this class reflects the abundance of the landcover type. Sensitivity to the *blueberries* parameter can be attributed to the dominance of this landcover type locally; Lonsdorf et al. (2009) concluded that the InVEST model was most sensitive to resources distributed at a small scale (Lonsdorf et al. 2009). The model also was sensitive to altering parameters for ground nesting bees, which accounted for 11 of the 14 modeled species.

Optimized model performance vs. expert opinion

Expert-informed parameterization is the typical approach for models used in conservation planning, and this approach was the baseline for comparison of the InVEST model for predicting pollinator abundance in wild blueberries. The informed and uninformed optimized models performed better than the expert-informed model, however, this does not invalidate the expert informed model. The best performing uninformed optimized model had parameter values that were very different than those values assigned through expert opinion. The informed, optimized model, which used expert survey derived parameter values that were then optimized based on the results of the sensitivity analysis, performed better than both the baseline expert-opinion model and the uninformed, optimized model. Although methods used to obtain expert opinion and synthesis of the results can affect the soundness of models parameterized with those results (Charney 2012), optimized models potentially overfit the data; the same dataset is used to calibrate and validate the model, and both the signal and the noise are fitted within the model. A more rigorous approach would include validation with an additional dataset as well as out-of-area model evaluation.

There are few examples of comparisons of expert opinion versus data driven model parameterization. Charney (2012) found that expert opinion assignment of model parameter values was unreliable for complex models requiring valuation of numerous parameters. The InVEST blueberry model required suitability rankings for 8 landcover classes, across three different seasons, and for two nesting guilds of bees. The InVEST model evaluating the Costa Rica coffee agroecosystem used expert assigned suitably rankings for 6 landcover classes and one floral season, and resulted in an $R^2 = 0.62$ (Lonsdorf et al. 2009). Although simplification of the model was appropriate for coffee, wild blueberries are a more complex crop system that is not adequately represented by a more simplified model.

Bee abundance based on landscape composition proportion vs. InVEST model predictions of bee abundance

In Maine's landscape, the proportion of both *deciduous/mixed forest* and *coniferous forest* are significantly and orthogonally correlated with the number of bees found within blueberry fields. The proportion of forest (deciduous and coniferous combined) surrounding Wisconsin apple orchards was similarly correlated with bee abundance, while the proportion of developed land surrounding a field was negatively correlated with bee abundance (Watson et al. 2011). The proportion of deciduous/mixed forest found within a 2000 meter buffer around a field may be a better predictor of bee abundance in the area immediately surrounding a blueberry field than the more complex InVEST model. The InVEST model predicts bee abundance across the landscape, while the simple proportional method provides predictions only within a blueberry field. I validated the blueberry InVEST model for only blueberry fields; however, bee abundance predictions in other landcover types were not evaluated with bee surveys. Although the InVEST model could be useful for large scale conservation planning, the simple proportional method is a useful tool for evaluating near farm pollinator habitat and bee abundance.

POTENTIAL LIMITATIONS

Although the InVEST model is a tool to examine relationships between land cover composition and bee abundance across a landscape, the tool has limitations. The biannual production cycle of lowbush blueberry, in which flowering fields during the fruiting year provide more floral resources than those fields in regrowth, introduces complexity into the InVEST model that was not incorporated into this application of the model. An additional limitation to my modeling efforts is that the field collected data spanned three years, while my landcover layer remained static through each model run. Thus, it did not capture any land use changes that could have occurred from the time it was created to the time the field sampling was conducted, as well as any interannual changes. Expert-informed parameter values that are inaccurate also potentially decrease model prediction performance.

Spatial models predict species distributions and abundances based on certain habitat conditions available across landscapes (Austin 2002, Guisan and Thuiller 2005, Elith and Leathwick 2009, Lonsdorf et al. 2011). Relationships between bees and land cover have been documented worldwide, and landscape scale predictive modeling, such as the InVEST Crop Pollination model, can use these relationships to predict bee abundance across the landscape (e.g., Kremen et al. 2004, Ricketts et al. 2008, Garibaldi et al. 2011). There are limitations to applying any tool, including those used to inform conservation efforts, and understanding the limitations is critical to ensuring appropriate use of the tool (Johnson and Gillingham, 2004). The InVEST model is sensitive to parameterization techniques used for applying the model to predict native bees in Maine's landscape. Additionally, more information is needed about bee abundances and species assemblages in Maine's different landcovers. Finally, a simpler, small scale model may be more appropriate than a complex, landscape scale model; understanding the purpose of the modeling effort and the desired outcome is a critical initial step in conducting a landscape assessment at an appropriate scale.

REFERENCES

Asher, J.S. and J. Pickering. 2013. Discover Life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila). http://www.discoverlife.org/mp/20q?guide=Apoidea_species.

Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modeling. Ecological Modeling 157(2-3):101-118.

Benjamin, F.E., J.R. Reilly, and R. Winfree. 2014. Pollinator body size mediates the scale at which land use drives crop pollination services. Journal of Applied Ecology 51(1): 1-10.

Beyer, H.L. 2012. Geospatial Modelling Environment (Version 0.6.0.0). (software). URL: http://www.spatialecology.com/gme.

Bushmann, S. 2013. Wild bee (Hymenoptera: Apoidea) communities associated with the lowbush blueberry agroecosystem of Maine. Dissertation, University of Maine, Orono, Maine.

Cane, J.H., and J.A. Payne. 1988. Foraging ecology of the bee Habropoda laboriosa (Hymenoptera: Anthophoridae), an oligolege of blueberries (Ericaceae: Vaccinium) in the southeastern USA. Annals of the Entomological Society of America 81: 419-427.

Cane, J.H. 1992. Soils of ground nesting bees (Hymenoptera: Apoidea): texture, moisture, cell depth and climate. Journal of the Kansas Entomological Society 64: 406 - 413.

Cane. J.H. 2001. Habitat fragmentation and native bees: a premature verdict? Ecology and Society 5(1): 3, [online] URL: http://www.consecol.org/vol5/iss1/art3/.

Charney, N.D. 2012. Evaluating expert opinion and spatial scale in an amphibian model. Ecological Modeling 242: 37 - 45.

Clarke, B., E. Fokoue, H.H. Zhang. 2009. Principles and theory for data mining and machine learning. Springer, New York, New York, USA.

Compton, B.W., K. McGarigal, S.A. Cushman, and L.R. Gamble. 2007. A resistant-kernal model of connectivity for amphibians that breed in vernal pools. Conservation Biology 21(3): 788 - 799.

Czembor, C.A., W. K. Morris, B.A. Wintle and P.A. Vesk. 2011. Quantifying variance components in ecological models based on expert opinion. Journal of Applied Ecology 48: 736-745.

Dibble, A., C. Stubbs, J. Ascher, and F. A. Drummond. unpublished data. Maine bee checklist.

Drummond, F.A. and C. Stubbs. 2003. Wild Bee Conservation for Wild Blueberry Fields. Extension Bulletin # 2111. University of Maine Cooperative Extension, Orono, Maine, USA.

Drummond, F.A. 2002. Honeybees and Lowbush Blueberry Pollination # 2079. University of Maine Cooperative Extension, Orono, Maine, USA.

Elith, J. and J.R. Leathwick. 2009. Species distribution models: Ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution and Systematics 40: 677-697.

Garibaldi, L.A., I. Steffan-Dewenter, C. Kremen, J.M. Morales, R. Bommarco, S.A. Cunningham, L.G. Carvalheiro, N.P. Chacoff, J.H. Dudenhoffer, S. S. Greenleaf, A. Holzschuh, R. Isaacs, K. Krewenka, Y. Mandelik, M. M. Mayfield, L.A. Morandin, S.G. Potts, T.H. Ricketts, H. Szentgyorgyi, B. F. Viana, C. Westphal, R. Winfree, and A.M. Klein. 2011. Stability of pollination services decreases with isolation from natural areas despite honey bee visits. Ecology Letters 14: 1062-1072.

Goffe, W.L., G.D. Ferrier, and J. Rogers. 1994. Global optimization of statistical functions with simulated annealing. Journal of Econometrics 60: 65 - 99.

Greenleaf, S., N. Williams, R. Winfree, C. Kremen. 2007. Bee foraging ranges and their relationship to body size. Oecologia 153: 589 - 596.

Guisan, A. and W. Thuiller. 2004. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8: 993 - 1009.

Henly, M. 2012. Re-establishing community boundaries in Downeast Maine: undersanding the roles of ethnicity, tenure of residence, economic & environmental conditions. Journal of Rural and Community Development 7(2):18-36.

Johnson, C.J. and M.P. Gillingham. 2004. Mapping uncertainty: sensitivity of wildlife habitat ratings to expert opinion. Journal of Applied Ecology 41 (1032 - 1041).

Kennedy, C.M. 2013. A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems. Ecology Letters 16 (5): 584 - 599.

Kirkpatrick, S., C.D. Gelatt, M.P. Vecchi. 1983. Optimization by simulated annealing. Science 220 (4598): 671 - 680.

Klein, A.M, B.E. Vaissiere, J.H. Cane, I. Steffan-Dewenter, S.A. Cunningham, C. Kremen, and T. Tscharntke. 2007. Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society 274: 303 – 313.

Kremen, C., N.M. Williams, R.L Bugg, J.P. Fay and R.W. Thorp. 2004. The area requirements of an ecosystem service: crop pollination by native bee communities in California. Ecology Letters 7: 1109-1119.

Kremen, C., N.M. Williams, M. A. Aizen, B. Gemmill-Herren, G. LeBuhn, R. Minckley, L. Packer, S. G. Potts, T. Roulston, I. Steffan-Dewenter, D. P. Vazquez, R. Winfree, L. Adams, E.E. Crone, S.S. Greenleaf, T.H. Keitt, A. Klein, J. Regetz, and T.H. Ricketts. 2007. Pollination and other ecosystem services produced by mobile organisms: a conceptual framework for the effects of land-use change. Ecology Letters 10: 299 - 314.

Lonsdorf, E., C. Kremen, T. Ricketts, R. Winfree, N. Williams and S. Greenleaf. 2009. Modeling pollination services across agricultural landscapes. Annals of Botany 103: 1589-1600.

Lonsdorf, E., T. Ricketts, C. Kremen, R. Winfree, S. Greenleaf, and N. Williams. 2011. Crop pollination services. Pages 168 – 187 *in* P.H. Kareiva, H. Tallis,. T.H. Ricketts, G.C. Daily and S. Polasky, editors. Natural Capital: Theory and Practice of Mapping Ecosystem Services. Oxford University Press - USA, New York, New York, USA.

Osgood, E.A. 1972. Soil characteristics of nesting sites of solitary bees associated with the low-bush blueberry in Maine. The Life Sciences and Agriculture Experiment Station, University of Maine at Orono: Technical Bulletin 59.

Martin, T.G., M.A. Burgman, F. Fidler, P. M. Kuhnert, S. Low-Choy, M. Mcbride, and K. Mengersen. 2011. Eliciting expert knowledge in conservation science. Conservation Biology 26 (1): 29 - 38.

McGarigal, K., SA Cushman, and E Ene. 2012. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. Available at the following web site: http://www.umass.edu/landeco/research/fragstats/fragstats.html

Michener, C.D. 1996. The bionomics of a primitively social bee, *Lasioglossum Versatum*. Journal of the Kansas Entomological Society 3: 193 - 217.

Michener, C.D. 2000. The Bees of the World.The John Hopkins University Press, Baltimore, Maryland, USA.

Millspaugh, J.J. and F. Richard Thompson. 2009. Models for planning wildlife conservation in large landscapes. Elsevier/Academic Press, Burlington, Massachussetts, USA.

Murray, J.V., A.W. Goldizen, R.A.O'Leary, C.A. McAlpine, H.P. Possignham, and S.L. Choy. 2009. How useful is expert opinion in predicting the distribution of a species within and beyond the region of expertise? A case study using brushtailed rock-wallabies *Petrogale penicillata*. Journal of Applied Ecology 46(4): 842-851.

Oliphant, T.E. 2007. Python for scientific computing. Computing in Science & Engineering 9 (90): 27 - 29.

Patricio-Roberto, G.B., and M.J.O. Campos. 2014. Aspects of Landscape and Pollinators - What is important to bee conservation? Diversity 6: 158 - 175.

Pettis, J.S. and Delaplane, K.S. 2010. Coordinated responses to honey bee decline in the USA. Apidologie 41(3): 256 - 263.

R Development Core Team. 2011. R: A language and environment for statistical computer. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, from http://www.R-project.org/.

Ricketts, T.H., J. Regetz, I. Steffan-Dewenter, S.A. Cunningham, C. Kremen, A. Bogdanski, B. Gemmill-Herren, S. S. Greenleaf, A. M. Klein, M.M. Mayfield, L.A. Morandin, A. Ochieng, and B.F. Viana. 2008. Landscape effects on crop pollination services: are there general patterns? Ecology Letters 11: 499-515.

Spear, S.F., N. Balkenhol, M. Fortin, B.H. Mcrae and K. Scribner. 2010. Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Molecular Ecology 19: 3576 - 3591.

Tepedino, A.J. 1979. The importance of bees and other insect pollinators in maintaining floral species composition. Great Basin Naturalist Memoirs 3: 139 – 150.

U.S. Census Bureau. (2014, January 06). *State & county Quickfacts: Hancock County, ME*. Retrieved January 31, 2014, from http://quickfacts.census.gov.

U.S. Census Bureau. (2014, January 06). *State & county Quickfacts: States, USA*. Retrieved January 31, 2014, from http://quickfacts.census.gov.

U.S. Census Bureau. (2014, January 06). *State & county Quickfacts: Washington County, ME*. Retrieved January 31, 2014, from http://quickfacts.census.gov.

Yarborough, D.E. 2009. Wild blueberry culture in Maine. Extension Bulletin # 2088. University of Maine Cooperative Extension, Orono, Maine, USA.

Yarborough, D.E. 2012. Statistics - Wild Blueberry Crop Statistics. University of Maine Cooperative Extension, Orono, Maine, USA. Retrieved April 9, 2014 from http://umaine.edu/blueberries/factsheets/statistics-2/statistics/.

Watson, J.C., A.T. Wolf, and J.S. Asher. 2011. Forested landscapes promote richness and abundance in native bees (Hymenoptera: Apoidea: Anthophila) in Wisconsin Apple Orchards. Environmental Entomology 40 (3): 621 – 632.

APPENDIX A: EXPERT OPINION SURVEY

READ ME FIRST

The first component of the survey is information regarding the landcover classes and bee species I am modeling. The page titled "Lookup Table - LANDCOVER", provides a look-up table with descriptions, and example floral resources for each of the 9 landcover classes. The page titled "Lookup Table - BEES", provides a look-up table with life history information on the bee species my modeling efforts are focused on.

The second component of the survey is where you come in. The page titled "Floral Resource Availability" and the page titled "Nesting Habitat" are set up to allow you write in a value from 1 (lowest quality) - 10 (highest quality) in each shaded cell. You will find more specific directions on what you are ranking, on the page titled "Floral Resource Availability" and the page titled "Nesting Habitat".

Table 9. Lookup Table - LANDCOVER.

This table provides a description of each landcover class and examples of potential bee forage plants/floral resources.				
IMPORTANT: The example floral resources listed below are listed to remind you of what is blooming at different times of the year. These lists do not imply abundance, nor due landcover classes with 7 examples imply better suitability for bees than landcover classes with 2 examples.				
Landcover Class	**Example: Feb. - April**	**Example: May - June**	**Example: July - September**	**Description**
Deciduous Forest edge (10 m)	*wild strawberry, willow*	*shadbush, raspberry, blueberry, blackberry, bunchberry, violet, bluebead lily, other spring herbaceous wildflowers*	*meadowsweet, pasture rose, asters, goldenrods*	This is the 10 meter strip on the edge of a deciduous or mixed forest patch
Developed	*dandelion, crocus, coltsfoot*	*azalea, chives, mints, apples, cherries*	*dandelion, oregano, bee balm, yarrow, roses, mints, goldenrods, asters*	This class represents all developed lands. Examples of this include rural, urban, suburban lands. This category does not include parks or developed, open spaces (see Landcover Class: Agriculture/Field).
Coniferous Forest	*trailing arbutus*	*sheep laurel, black huckleberry, wintergreen*	*raspberry, blackberry, goldenrod, aster*	This class represents coniferous forest, including regenerating forest and the edge
Coniferous Forest - clearcut	*dandelion, red maple, trailing arbutus*	*sheep laurel, blueberry, black huckleberry*	*raspberry, blackberry, goldenrod, aster*	This class represents clearcut or recently harvested coniferous forest, including the edge
Deciduous/Mixed Forest	*maple, willow*	*oak, columbine, honeysuckles, shadbush, viburnum, other spring herbaceous wildflowers*	*meadowsweet, aster*	This class represents all deciduous and mixed forest
Emergent/Scrub Shrub Wetlands	*willow, red maple*	*leatherleaf, rhodora, cranberries, violets*	*St.John's wort, meadowsweet, steeplebush, summersweet, aster, shrubby cinquefoil*	This class includes both emergent and scrub-shrub wetlands
Wetlands/Water	*willow, red maple*	*highbush blueberry, mountain holly*	*pickerelweed, water lillies, purple loosestrife*	This class represents all other wetlands (marine, riverine, and estuarine) and open water
Agriculture/Field	*dandelion, willow*	*alfalfa, clover, hawkweed*	*vegetable crops, goldenrods, asters, meadowsweet*	This class represents cultivated crops (except blueberries), pastures, grasslands and developed open space (i.e., parks)
Blueberries	*willow*	*blueberry, bunchberry, violet, sheep laurel*	*vetch, St.John's wort, butter and eggs, goldenrods, asters, dogbane*	This class represents both wild blueberries and managed blueberry fields

Table 10. Lookup Table - BEES.

This table provides life history information on the bee species we are using in our modeling efforts.

Species	Typical Foraging Distance (m)	Nest Substrate	General Flight Season
Andrena carlini	598	ground	March - August
Andrena carolina	246	ground	April - July
Andrena vicina	569	ground	March - August
Augochlorella aurata	60	ground	April - October
Colletes inaequalis	1091	ground	March - July / August - September
Halictus ligatus	148	ground	throughout the year
Lasioglossum acuminatum	186	ground	April - October
Lasioglossum cressonii	63	cavity	March - October
Lasioglossum heterognathum	16	ground	April - September
Lasioglossum leucocomum	31	ground	March - October
Lasioglossum pectorale	81	ground	March - November
Lasioglossum versatum	79	ground	March - October
Osmia atriventris	186	cavity	April -July
Osmia inspergens	495	cavity	May - June
Queen - Bombus ternarius	5767	ground and cavity	April - October
Queen - Bombus vagans	4415	ground and cavity	May - October
Queen - Bombus spp.	7554	ground and cavity	Feb - November
Worker - Bombus ternarius	966	ground and cavity	June - October
Worker - Bombus vagans	1261	ground and cavity	June - October
Worker - Bombus spp.	2125	ground and cavity	June - November

Table 11. Floral Resource Availability

First, let's think about floral resources (forage for bees) in the landscape, across the seasons:						
This is a ranking based on the relative abundance of floral resources/flowering plants in each landcover class throughout the seasons. Starting in the column titled "February - April" set the landcover class with the greatest availability of floral resources during February - April, to 10, and give all other landcover classes that column a value relative to this maximum value (between 1 - 10). Repeat this exercise for the column titled "May - June (blueberry bloom)", "July - September", and "Yearround (February - September)". See page titled "1. Lookup Table - LANDCOVER" for a description and examples of potential bee forage within each landcover class during the different months. To the right I have provided an example scoring in the column titled "EXAMPLE" and reasoning for my scoring in the column titled "REASONING". You do not need to provide your reasoning, I just wanted to demonstrate why I assigned the values I did. It is okay to leave a cell blank if you are unsure.					**EXAMPLE**	**REASONING**
Landcover Class	**February - April**	**May - June (blueberry bloom)**	**July - September**	**Yearround (February - September)**	**February - April**	**Available forage**
Deciduous Forest edge (10 m)					6	willows, wild strawberries
Developed					10	crocuses, dandelions, coltsfoot
Coniferous Forest					2	trailing arbutus
Coniferous Forest - Clearcut					-	unsure
Deciduous/Mixed Forest					9	maples, willows
Emergent Wetlands/Scrub Shrub					8	willow, red maple
Wetlands/Water					2	not much flowering
Agriculture/Field					6	possibly apples, choke cherries, dandelions
Blueberries					1	blueberry isn't flowering yet

Table 12. Nesting Habitat

Next, let's switch gears and think about nesting habitat based on the landcover classes, and soil types:

LANDCOVER CLASS

Native bees are known to nest in both the ground, and in cavities/rotten wood and stems. This is a ranking of the availability of nesting for native bees within a given landcover class. Starting in the column titled "Ground Nester", set the landcover class with the greatest availability of nesting habitat for ground nesters to 10, and give all other landcover classes a value relative to this maximum value (between 1 - 10). Repeat this exercise for the column titled "Cavity Nester". It is okay to leave a cell blank if you are unsure. Ground nesters include bees that nest in the soil, and cavity nesters in rotten wood, cavities and stems. See the page titled "2. Lookup Table - BEES" for a list of species in each category and information on their life history.

Landcover Class	Ground Nester	Cavity Nester
Deciduous/Mixed Forest edge (10 m)		
Developed		
Coniferous		
Deciduous/Mixed Forest		
Emergent Wetlands/Scrub Shrub		
Wetlands/Water		
Agriculture/Field		
Blueberries		

SOIL TYPE

For those native bees that nest in the ground, please rank the soil types based on the potential availability of nesting habitat. For the column titled "Ground Nester", set the soil type with the greatest availability of nesting habitat to 10, and give all other soil types a value relative to this maximum value (between 0 - 10). It is okay to leave a cell blank if you are unsure.

Soil Type	Ground Nester
coarse, sandy, well drained soil	
coarse, sandy, poorly drained soil	
sandy - loam, well drained soil	
sandy - loam, poorly drained soil	
silty - clay, well drained soil	
silty - clay, poorly drained soil	

APPENDIX B. RESULTS OF ADDITIONAL INVEST TESTS

I investigated relationships between the landscape and native bee abundance using the InVEST Crop Pollination Model adapted to Downeast Maine's landscape. I compared model performance using different spatial landcover data layers, modeled extents and validation datasets, with lowbush blueberry fields in Downeast Maine as the focal study system. Much of the methods are described above, but below I describe information associated with a few of the additional runs that I conducted.

Description of Study Extents and Field Sites

Maine produces the greatest area (>24,000 ha) of managed, lowbush blueberries of any state (Yarborough 2009). Most of this management activity is in Downeast Maine, in Hancock and Washington counties. We evaluated the InVEST model for three extents (Figure 3.) spanning Downeast Maine, reflecting differences in landcover type, validation datasets, and patch size across this region.

I evaluated the predictive ability of the InVEST model across three spatial extents; the first extent (Eastern) covers 3000 km^2 of the region (Figure 3.). Eight focal blueberry fields (< 1 - 11 hectares) are located within this extent. The second extent (Blue Hill; Figure 3.) covers 705 km^2 of southwestern Hancock County, and includes 26 focal blueberry fields (< 1 - 17 hectares). There are 40 focal blueberry fields (< 1 - 17 hectares) in the third extent (Eastern), which spans 4,802 km^2 of the blueberry growing region.

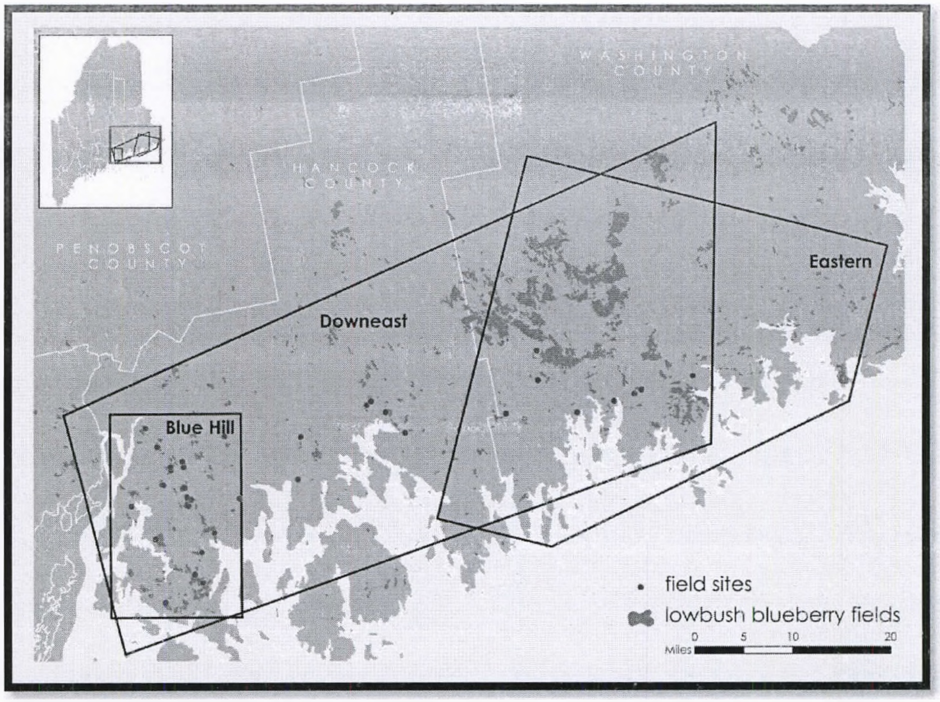

Figure 3. Extents modeled and blueberry field sites used for validation of additional InVEST model runs, Maine, USA.

Methods

Landcover layer used

In addition to the methods described above, I also updated the landcover layer with satellite imagery that I classified. I purchased a single 10-m hyper spectral SPOT image of a 3,600 km^2 area of Washington County from May 2011 in an attempt to update the blueberry field coverage within the Eastern extent only (Airbus Defence and Space 2014; Figure 3.). To improve the classification among landcover types, I used the MELCD as a guide to extract all pixels from the image that were not classified as water and wetlands and then conducted an isocluster unsupervised classification on them (ArcGIS ® version 10.0; Environmental Systems Research Inc., Redlands, CA, United States). Following the unsupervised classification, I developed training sets for landcover classes that were grouped with the *blueberries* class in the results of the unsupervised classification. Training sets were developed for roads and gravel pits, conifers and blueberry fields using the MELCD dataset and aerial imagery (Bing Maps 2010). These training pixels were used in a maximum likelihood supervised classification and the subsequent classification that represented blueberries was added to the MELCD *blueberries* class.

The final landcover dataset included 42 classes reclassified into 8 landcover types: *deciduous/mixed forest edge*, *developed/other*, *coniferous forest*, *deciduous/mixed forest*, *emergent/shrub-shrub wetlands*, *other wetlands/water*, *agriculture/field* and *blueberries* (Table 1). Although the resolution of the final updated dataset remained at 5-m, we conducted the InVEST analysis at a 10-m resolution to decrease analysis time.

Landcover pattern description

I compared landscape pattern metrics for the three modeled extents (Figure 3.) with Fragstats 4.2 (McGarigal et al. 2012). For each landcover class I calculated the proportion of the extent in that class, patch density (number per 100 hectares (ha)), mean patch area (ha), and a measure of spatial configuration (i.e., interspersion/juxtaposition index). I also calculated a landscape scale mean patch area (ha) and interspersion / juxtaposition index (IJI) for each model extent.

Results

Pattern metrics

More than half of the region bounded by the Eastern extent was *coniferous forest* (24.5%) and *wetlands/water* (27.5%), and the mean patch sizes of both the *deciduous/mixed forest* and *blueberries* classes were larger than the mean patch area of the entire landscape (Table 13.; Table 14.). The Eastern landscape IJI was 73.7 (Table 14.). *Coniferous forest* and *deciduous/mixed forest* class comprised more than half of the landscape in the Blue Hill extent. The mean patch area for each class and the landcover mean patch area were similar, with the exception of the *coniferous forest* mean patch area, which exceeded all other patch sizes. The landscape IJI was 74.12. Additionally, the *coniferous forest* and *deciduous/mixed forest* classes made up over half of the landscape in the Downeast extent, and the landscape IJI was 73.6.

Table 13. Proportion of land and mean patch area (ha) per class for each extent

Class	Eastern		Blue Hill		Downeast	
	% land	mean patch area (ha)	% land	mean patch area (ha)	% land	mean patch area (ha)
Deciduous/Mixed Forest edge	4.0	1.5	4.3	1.2	4.3	1.3
Developed/Other	1.5	1.1	4.4	2.2	2.7	1.8
Coniferous Forest	24.5	6.6	34.4	10.0	28.7	8.8
Deciduous/Mixed Forest	24.1	11.6	21.3	6.0	26.3	10.3
Emergent/Scrub-Shrub Wetland	11.9	4.4	8.6	3.1	10.6	3.7
Wetlands/Water	27.5	5.9	20.0	5.9	21.0	5.3
Agriculture/Fields	1.3	1.0	3.3	1.6	1.9	1.3
Blueberries	5.3	11.0	3.7	5.6	4.4	10.1

Table 14. Mean patch area (ha) and interspersion-juxtaposition index (IJI) for each extent

	Extent		
	Eastern	Blue Hill	Downeast
mean patch area (ha)	5.3	4.6	5.2
IJI	73.7	74.1	73.6

Model prediction and correlations

Combining the SPOT image based blueberry classification with the MELCD landcover did not result in different InVEST model predictions of bee abundance for the Downeast extent, however, total bee abundance was significantly correlated with the InVEST model bee abundance estimate for both the SPOT-enhanced and non-enhanced landcover when the modeled bee species were restricted to those with estimated foraging distances < 200 m (9 bee species, Pearson's $r = 0.77$; $P = 0.02$) (Table 15.). Correlation of the model-predicted and sampled bee abundance increased with restriction of the bee species that have an estimated foraging range < 100 m (6 bee species, Pearson's $r = 0.86$; $P < 0.01$).

InVEST predicted and sampled bee abundance were not significantly correlated in the Blue Hill extent, regardless of grouping by foraging distance or the number of bee species included.

I observed significant correlations when modeling both 14 species communities (Pearson's $r = 0.32$; $P = 0.04$) and bees that forage < 200 m (9 bee species, Pearson's $r = 0.36$; $P = 0.02$) for the Downeast extent. A non-significant trend similar to correlations observed from previously described model runs was observed when modeling bee species that forage < 100 m (6 bee species, Pearson's $r = 0.26$; $P = 0.08$).

Table 15. Pearson's r correlation and P values between InVEST model-predicted and observed bee abundance for the three focal spatial extents in Maine.

Extent	Landcover	Species Modeled	r	P
Eastern	updated with SPOT	14 species	0.52	0.19
		14 species	0.52	0.19
	updated with SPOT	9 species (foraging < 200 m)	0.77	0.02
		9 species (foraging < 200 m)	0.77	0.02
		6 species (foraging < 100 m)	0.86	0.01
Blue Hill		14 species	0.32	0.12
		9 species (foraging < 200 m)	0.33	0.11
Downeast		14 species	0.32	0.04
		9 species (foraging < 200 m)	0.36	0.02
		6 species (foraging < 100 m)	0.26	0.08

Discussion

Spatial landcover dataset, species and extent modeled effects on model output

The relationship between the InVEST Crop Pollination model's predictions and observed native bee abundance in Maine's landscape did not vary depending on the spatial dataset used, but did depending on both the species and extent modeled.

The addition of the SPOT updated *blueberries* class did not alter the explanatory power of the InVEST model across the Eastern extent. This was encouraging; large differences between the results would have required me to update the *blueberries* class through the purchase of additional SPOT imagery, increasing project expenses.

There was a difference with significance and prediction power within all extents when I changed the number of species modeled. It is not surprising that results ranged from significant to non-significant across the Eastern extent modeling efforts; this could be due to my small sample size of 8 field sites. Overall, correlation between observed and predicted abundances was best within the Eastern extent when modeling only 6 species (foraging distance < 100 m). The major landcover class within a 100 meter foraging buffer around the study sites located in the Eastern extent is *blueberries*. Previous work with the InVEST model has indicated that model predictions are most sensitive to the floral resources provided at the smaller scale (Lonsdorf et al. 2009). Additionally, it has been suggested that smaller bees (such as those foraging < 100m) are more strongly influenced by local, field scale resources (Benjamin et al. 2014).

Similarly, the fact that that the moderately positive correlations for the Blue Hill extent were non-significant could also be due to small sample size (26 field sites). The patch size of the local resources (*blueberries*) are much smaller than the two other extents, with a mean patch area of 5.6 ha (Table 14.), compared to mean *blueberries* patches of 11.0 ha and 10.1 for the Eastern and Downeast extents respectively. Additionally, the landscape is quite different within the Blue Hill extent than it is in Eastern or Downeast extents. Specifically, there is both a greater proportion of and larger patches of coniferous forest within the Downeast extent (Table 13.). The smaller patch size present within the Blue Hill extent could limit the predictive power of the InVEST model. Smaller patches of resources may not be adequately reflected in the spatial landcover layer used; this was suggested as the reason that the InVEST model did not accurately predict bee abundance in other landscapes (Lonsdorf et al. 2009).

The weak to moderately weak positive relationships observed between predicted and observed abundances across the Downeast extent varied little when the number of species modeled was altered. This was encouraging as this modeling effort spanned much of the blueberry growing region and included all of the validation datasets.

BIOGRAPHY OF THE AUTHOR

Shannon Chapin was born in Lock Haven, Pennsylvania in 1984, and graduated from Central Mountain High School in Mill Hall, Pennsylvania in 2003. Shannon earned a Bachelors of Science in Geography, with minors in Wildlife and Fisheries Sciences, and Climatology from The Pennsylvania State University in University Park, Pennsylvania in 2007. She received a post-baccalaureate certificate in 2010 in Geospatial Sciences from Humboldt State University in Arcata, CA. Prior to returning to graduate school, Shannon worked for 5 years for various federal agencies, universities and a private firm as a field ecologist and GIS Analyst. After receiving her degree, Shannon plans to continue her career as a Geospatial Analyst at the Southern Environmental Law Center, a non-profit environmental advocacy organization, in Chapel Hill, North Carolina.

Shannon is a candidate for the Master of Science degree in Ecology and Environmental Science from the University of Maine in May 2014.