SW: Hi! I’m Sadie Witkowski

IM: And I’m Ian Martin

SW: And you’re listening to Carry the Two, a podcast from the Institute for Mathematical and Statistical Innovation

IM: Aka- IMSI! This is usually the podcast where Sadie and I talk about how math and statistics impact the world around us.

SW: “Usually,” because today we’re doing something a bit different and unique!

TC: Hi/hello/hey!

SW: While we’re in between our more in-depth seasons, we wanted to bring you something a little different. So for this 6-episode run between our main seasons, we have a special guest joining us.

IM: Would you like to introduce yourself?

TC: Sure, I’m Tiffany Christian. I’m a rising fourth year PhD student at Northwestern University in the statistics department and I used to know how to unicycle.

SW: That's something you can forget?

TC: You know I haven't tried since I learned so potentially you can forget how to unicycle.

IM: Do you just have like really good balance? Like how do you even get up there?

SW: I bet you had a six-pack from using your core muscles

TC: Oh yeah, 13-year-old me was ripped.

SW: Ok, but to get back on topic, Tiffany is joining us as our Statistician-in-Residence to help us tackle statistics in the world of environmental and climate science.

IM: Yay! I love giving my environmental nightmares statistical backing!

SW: Alright, let’s get this show on the road!

[Music Break]

TC: I know you guys said that we would be focused on climate change, but I think today’s topic is a parallel issue that’s certainly affected by climate change.

SW: That’s right, today we’re covering invasive species. Which, as a reminder to all our listeners out there, is when a plant or animal or other creature leaves its original habitat.

IM: It’s basically like how gardners define weeds as just something they don’t want growing in their garden.

SW: Yeah, so the USDA (the us dept of agriculture) defines them as a non native creature that comes into an ecosystem. That whose introduction causes or is likely to cause either economic and or environmental harm. So invasive species can be introduced into an area by anything from, ships’ ballast water, to logging for firewood that brings along insects, to accidental release, sometimes by people.

IM: I have a very good example of that. And by good, I mean horrifying. So I was listening to this Bill Bryson book about Australia called 'In a Sunburned Country'. And he mentioned that there was a British dude named Thomas Austin, who in 1859 really wanted to hunt rabbits for sport while in Australia. And so he had 13 rabbits sent to him. Because you know, god forbid you find a native animal to hunt. Or pick a different pastime. Anyway, it only took 50 years for the rabbits to take over the whole country. It had devastating effects on the vegetation because they overgrazed and exposed the topsoil, which then blew away. This was followed unfortunately by a series of droughts. So now there are many places in the country where you can't plant vegetation because all the nutrients have gone from the soil.

TC: Oh yeah, these are the kinds of horror stories that ecologists are trying to avoid.

SW: But at least with the rabbits, they’re an easy invasive species to spot. Even if they aren’t so easy to remove. But some invasive species are much smaller and harder to notice. It can be really difficult to know which areas are particularly faced with a threat and that need extra protection. And so that's what we're going to be talking about today, using this really cool study that uses a spatial analysis to try to recognize and protect specific areas of a forest.

TC: That’s right, today I want to talk about this 2014 study by Liang and colleagues, titled Spatial analysis facilitates invasive species risk assessment.

IM: Ok, so what's the creature we're talking about here?

SW: In this case, the authors were trying to look at different regions and regional hotspots for what is called the hemlock, wooly adelgid. It's an invasive insect that's threatening native forests and particularly preys on Hemlock and spruce trees. It was introduced as an invasive species. And now it's starting to take over these forests.

And so the study Tiffany mentioned is trying to see the areas that are most at risk, where we should be spending our money and our time for preventative measures to be the most cost effective. Basically, the study is trying to look at ways to figure out the probability of invasion based on, for example, known locations of where the bug has been seen in the past or other environmental factors. And to find a way to reduce the uncertainty in this modeling so that they can really be like, yes, these are the top five hotspots where we need to intervene. And these are five areas we should keep an eye on, but are probably going to be okay.

IM: I get the reasoning behind the study, but I know nothing about spatial statistics. Tiffany? Help a guy out?

TC: Yeah. So, spatial statistics is an area of statistics where the data's location is just as important as the data itself.

whatever variable we are interested in, its important that we found this particular thing happening in this particular area. Location provides a lot of information in this paper. You can imagine, if you pick a point at random in a forest there are many different environmental variables you could look at: elevation, temperature, soil conditions, sun exposure, how windy it is, and the list goes on.

Since they wanted to investigate areas that were vulnerable to hemlock wooly adelgid, looking at stands of hemlock and spruce in the mountains of Kentucky. They were particularly interested in modeling the hemlock wooly adelgid areas that would be hard to monitor consistently, so they picked mountainous, or high elevation, areas that would be difficult to access.

They wanted to have a way to know, like, "Oh, these are some vulnerable populations, how often do we need to send out people to monitor how what factors are going to influence their vulnerability". One of the things about the location that was important, or that they're the trees that the insect attacks are there. So that's one spatial variable.

IM: Must have trees that are under threat, check!

TC: They also wanted to know like, okay, what are some of the pathways for the insect to come into this particular area?

So they also were looking at old railroad routes and active recreational trails, like hiking trails, which create corridors through these natural spaces.

IM: Ye olde timey hitchhikers catching a ride? Cuz insects are famously avid hikers and love taking the train. Got it.

SW: Do the hemlock wooly adelgids get their own amtrak stop? [Laugh]

TC: Yeah, literally hitch a ride. So they had the railroad routes, the incline of the terrain, the temperature, different climate variables, and wind because it's a small insect, so they could be blown.

So they had all these variables, and they basically combined that information with data of where the Adelgid had already been found, and data that showed the adelgid might wanna live here.

IM: And these are all spatial variables because they differ across the map. Yes?

SW: Right, although some of this data is a bit incomplete. Like, they know where the adelgid has been found so far, but they're not testing everywhere necessarily.

TC: Yeah, they had like, presence data of, okay, we spotted it here and here. They didn't have absence data, because it's a small thing to check for. So just because we didn't see it in a place doesn't mean that it's not there.

And they use a particular kind of statistical model where that kind of data where you only have presences, it's okay to use. And so, they put all this into a model.

SW: And just so our listeners know, we don't mean a physical model, like a miniature forest. Rather, in this case we're talking about a computer program containing lots of mathematical equations that help us make predictions about the system's behavior. Think of something like a weather prediction model.

IM: So what does the model spit out?

TC: Basically, the model tells us which areas are susceptible. So like, this particular mountainous area is highly susceptible, but this particular mountainous area is not very susceptible.

And they use a couple different models to do this, with several different amounts of high susceptibility areas. And so at that point, they could compare models and say, that one that had a lot of highly susceptible areas, that's like, Okay, we have to protect everything.

SW:So it could not be feasible to do that.

TC: Right. But also could mean like, well, we need more resources from the state or whoever to try to get more hands or feet on the ground for that.

Or we could go with the smaller, predicted susceptible areas, and really concentrate our resources on that smaller area that was predicted by the model.

So they could use this model comparison to determine where the intervention was needed most.

They also did a thing where based on the results of the models, they were able to look at areas that were highly susceptible and see like, is there a hotspot. So a hotspot is just a way of saying like, are all of your neighbors sharing a similar value? So if all your neighbors are all highly susceptible, you might be in a highly susceptible hotspot.

IM: So if I'm in a theater and I can hear six to ten people around me with a covid cough, I'm in a highly susceptible hotspot.

SW: [laugh] yeah, same idea.

IM: Also, just listening to you list all of the variables. It's just another reminder of how complicated everything is surrounding environments and ecosystems. And just how delicate it is. Like, how one action can have consequences that we don't even think about.

SW: I'm still caught by the complexity of the data. For the adelgid, you have a presence signal, not an absence signal. But then you have variation of elevation, which is going to give you multiple degrees that you can be looking at. And then you have railroad presence or absence, which is a binary. It makes me think of those like transparencies that you could layer over a map, on top of each other. But some are binary and some aren't and how complicated that is to try to model. All that work, just to come up with a value where you can just tell some policy maker, “I did the math and we need to worry about block 3B.” Like, that's crazy.

TC: Yeah, the transparency is a really good analogy, because that’s the software that these kinds of researchers use. The software is called GIS, and it's literally layers. You're literally stacking layers of data on top of each other. And you could pick from the layers, which are going into your model. So you can remove things and add things based on this whole surface of data that you have. So that's how I like to think about spatial statistics is that these things are happening on the surface, because they literally are. They're happening on the surface of, well, in this case, the earth. So at each point, there's so many different variables we could measure at any given time. And everything is really connected.

IM: And some are connected by railroads! But actually, you mentioned that the authors of the paper seemed surprised by how important the railroads were.

TC: The researchers were surprised how little wind was important in their model, given that in other papers it is a known dispersal pathway for the adelgid. They think for their particular areas of interest, the patches of adelgid habitat are more dispersed. So the wind could blow adelgids into an area where they don’t have a good chance of surviving with less hemlock trees. So because of the patchiness of suitable habitat, things that create corridors between the hemlock tree patches proved to be really important in their models. Where the wind will indiscriminately blow a bug, the corridors act like more of a highway between patches, where the bugs can move on birds, deer, humans, and other wildlife.

IM: Wow. I'm just sort of reveling in the complexity of it all.

SW: Yeah, it's just really cool to think about, like trying to make these really targeted preventative measures, based off of having a really good understanding of all these diverse variables in the ecosystem. I just think that's so cool. And a good way to think about improving conservation, like getting your best bang for buck on conservation.

IM: Statistics to the rescue!

[music plays]

[AD BREAK]

SW: As always, don't forget to check out our show notes in the podcast description for more about this research on invasive species.

IM: And if you like the podcast, subscribe to the show and give us a review on apple podcast or spotify or where ever you listen.

TC: By rating and reviewing the show, you really help spread the word about Carry the Two so that other listeners can discover the show.

SW: And for more on the math research being shared at IMSI, be sure to check us out online at our homepage: IMSI dot institute. We're also on twitter at IMSI underscore institute, as well as instagram at IMSI dot institute!

IM: How do you spell that again?

SW: That's IMSI, spelled I M S I.

TC: And what are your mathy questions? Maybe you have an idea for a story on how mathematics and statistics connect with the world around us.

IM: Send us an email with your idea!

SW: You can send your feedback, ideas, and more to sadiewit AT IMSI dot institute. That's S A D I E W I T at I M S I dot institute.

IM: We'd also like to thank our audio engineer, Tyler Damme for his production on the show.

TC: And music is from Blue Dot Sessions.

SW: Lastly, Carry the Two is made possible by the Institute for Mathematical and Statistical Innovation, located on the gorgeous campus of the University of Chicago. We are supported by the National Science Foundation and the University of Chicago.

Until next time!

IM/TC: See ya!