Science of the Total Environment 630 (2018) 578–586

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Predicting risk of trace element pollution from municipal roads using site-specific soil samples and remotely sensed data

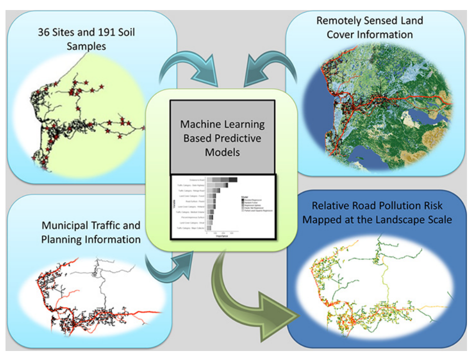
Mari Kathryn Reeves [a,*], Margaret Perdue [b], Lee Ann Munk [c], Birgit Hagedorn [c,d]

[a] *Strategic Habitat Conservation Program, Pacific Islands Fish and Wildlife Office, United States Fish and Wildlife Service, Honolulu, HI, USA*
[b] *Water Resources Branch, National Wildlife Refuge System, United States Fish and Wildlife Service, Anchorage, AK, USA*
[c] *University of Alaska, Anchorage, Department of Geological Sciences, Anchorage, AK, USA*
[d] *University of Washington, Quaternary Research Center, Seattle, WA, USA*

HIGHLIGHTS

- Describes trace element pollution risk from roads, given traffic and land cover category.
- Used predictive analytics to prioritize road maintenance by environmental pollution risk.
- Spatially-explicit model predictions used to map roads needing maintenance.
- Multi-model machine learning approach predicted distribution of trace elements in soils.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 11 October 2017
Received in revised form 1 February 2018
Accepted 14 February 2018
Available online xxxx

Editor: Simon Pollard

Keywords:
Random forests
Boosted regression
Elastic net regression
General linear models
National land cover dataset
Trace elements

ABSTRACT

Studies of environmental processes exhibit spatial variation within data sets. The ability to derive predictions of risk from field data is a critical path forward in understanding the data and applying the information to land and resource management. Thanks to recent advances in predictive modeling, open source software, and computing, the power to do this is within grasp. This article provides an example of how we predicted relative trace element pollution risk from roads across a region by combining site specific trace element data in soils with regional land cover and planning information in a predictive model framework. In the Kenai Peninsula of Alaska, we sampled 36 sites (191 soil samples) adjacent to roads for trace elements. We then combined this site specific data with freely-available land cover and urban planning data to derive a predictive model of landscape scale environmental risk. We used six different model algorithms to analyze the dataset, comparing these in terms of their predictive abilities and the variables identified as important. Based on comparable predictive abilities (mean R^2 from 30 to 35% and mean root mean square error from 65 to 68%), we averaged all six model outputs to predict relative levels of trace element deposition in soils—given the road surface, traffic volume, sample distance from the road, land cover category, and impervious surface percentage. Mapped predictions of environmental risk from toxic trace element pollution can show land managers and transportation planners where to prioritize road renewal or maintenance by each road segment's relative environmental and human health risk.

Published by Elsevier B.V.

* Corresponding author at: Strategic Habitat Conservation Program, Pacific Islands Fish and Wildlife Office, United States Fish and Wildlife Service, 300 Ala Moana Blvd, Honolulu, HI 96850, USA.
E-mail address: mari_reeves@fws.gov (M.K. Reeves).

https://doi.org/10.1016/j.scitotenv.2018.02.171
0048-9697/Published by Elsevier B.V.

1. Introduction

Roads cause ecological impacts due to soil and water pollution (Bäckström et al., 2003; Dorchin and Shanas, 2010; Krein and Schorer, 2000; Mangani et al., 2005; Sutherland and Tolosa, 2001), wildlife mortality from roadkill (Gibbs and Shriver, 2005; Lode, 2000; Mazerolle, 2004) and their role as barriers to organismal movement across ecosystems, leading to habitat fragmentation (Forman and Deblinger, 2000; Lode, 2000; Trombulak and Frissell, 2000). In addition, roads and associated airborne pollutants may harm human health (Kampa and Castanas, 2008). Factors driving the amount and type of road pollution include traffic volume (Kayhanian et al., 2002), road surface materials (Edvardsson and Magnusson, 2009; Norman and Johansson, 2006), driving conditions such as road grade and need for acceleration or braking (Egodawatta and Goonetilleke, 2008; Ragione and Giovanni, 2016), climate (Memon and Butler, 2005) and activities associated with climate like use of studded tires (Hussein et al., 2008), traction sand (Kupiainen et al., 2003), and salt (Corsi et al., 2010; Novotny et al., 2008). Characteristics of adjacent environments like galvanized steel structures (e.g., bridges or guardrails) or agricultural land use may contribute trace elements to road corridors (Blok, 2005; Wei and Yang, 2010). Trace elements originating from traffic exhaust or from the road bed itself may migrate from the roadway into adjacent air, water, and soils, potentially contaminating areas adjacent to roadways with levels of trace elements exceeding human health or ecotoxicological thresholds.

Road maintenance options to minimize risk from road-based trace element pollution include reducing road surface wear (e.g., by paving (Kupiainen et al., 2003) or applying dust suppressants (Edvardsson and Magnusson, 2009), removing dust by sweeping or washing streets (Vaze and Chiew, 2002; Westerlund and Viklander, 2006), slowing traffic (Hussein et al., 2008; Williams et al., 2008), or engineering the road bed or shoulder to retain pollutants in adjacent soils (Piguet et al., 2008). Most municipal entities responsible for road maintenance have limited budgets, so an ability to analyze and predict where road maintenance is most needed to protect the environment and human health from trace element pollution will help municipalities prioritize road maintenance in areas where these activities have the greatest social and environmental benefit.

Another core question for land managers considering new road construction projects in remote areas is whether paved or gravel roads add more trace element pollution to adjacent environments; but there is disagreement in the literature on this topic (Claiborn et al., 1995; Hussein et al., 2008; Kupiainen et al., 2003; Williams et al., 2008). The basis of disparate findings in different studies probably reflects environmental heterogeneity and complexity in the interactions among factors causing the trace element deposition. For example, paved roads often carry more vehicles moving at higher speeds than unpaved roads—making it hard to disentangle the individual effects of road surface, road traffic, and vehicle velocity. Moreover, different land use in adjacent areas is likely to influence the sources of pollutants and the width of the trace element deposition zone (Pocock and Lawrence, 2005).

Statistical models are integral to modern understanding, prediction, and management of complex phenomena. Current modeling techniques are improving predictions of outcomes driven by interrelated and nonlinear variables. The field of machine learning is a branch of computer science and computational statistics that uses a variety of different mathematical or grouping algorithms to generate predictions from data. Machine learning techniques can solve regression (numerical prediction) or classification (binomial or multinomial outcome) problems. When the user provides the model already classified training data, this is considered supervised machine learning. Though there has been substantial work modeling road pollution chemistry (Lin et al., 2008) and contaminant fate and transport (Murakami et al., 2004; Omstedt et al., 2005; Ragione and Giovanni, 2016), to our knowledge no study has used supervised machine learning based predictive analytics to better understand and manage road pollution risk, particularly with models that explicitly account for space in the predictions.

Here we present such an analysis. To perform it, we acquired site specific soil chemistry measurements of trace element pollutants adjacent to paved and gravel roads on the Kenai Peninsula of Alaska. We then combined these with regional transportation planning information available for this road network and two different remotely sensed national land cover datasets (land cover category and percent impervious surface). We hypothesized that more traffic and impervious surface around the sample would increase soil contamination and that land cover at the point of the sample would affect trace element concentrations in soil, for example dense forest cover may filter trace elements from auto exhaust or developed land cover may contribute trace elements to the road corridor. We compared six different model frameworks to understand these relationships and averaged their outputs to predict trace element deposition at over 50,000 randomly placed points within 80 m of mapped roads across this landscape. Ultimately these methods allowed us to make regional maps depicting relative trace element pollution risk by each road segment in the Kenai Peninsula Borough and provide the predictions as a data file so that land managers can summarize the information themselves as needed.

2. Materials and methods

2.1. Site selection

The primary study goals were to understand the effects of roads on the incidence of malformed amphibians in nearby sensitive habitats (Hayden et al., 2015; Reeves et al., 2011, 2010, 2008), so 36 sites on the Kenai Peninsula of Alaska were chosen using a stratified random sampling design. The Kenai Peninsula Borough of Alaska is a 6.4 million ha area (Fig. 1) with a human population of approximately 55,400. We stratified first on wetland presence (to meet other study objectives), then on road surface (paved or gravel) because metal fate and transport mechanisms may differ between gravel and paved roads. We also divided the study area into 5 spatial sectors to ensure interspersion of sites on the landscape. These sectors varied in road density and traffic as well as land use and degree of urbanization.

Within the chosen area (Fig. 1), we included all roads classified as a street or main road on the Kenai Peninsula Borough's (KPB) GIS "Roads" layer in 2010 (available: http://www.kpb.us/gis-dept/kpb-data-downloads/transportation). We created a 1 km buffer around the streets and main roads in this layer (which has since been updated to include more detail about road categories) and chose possible sites only

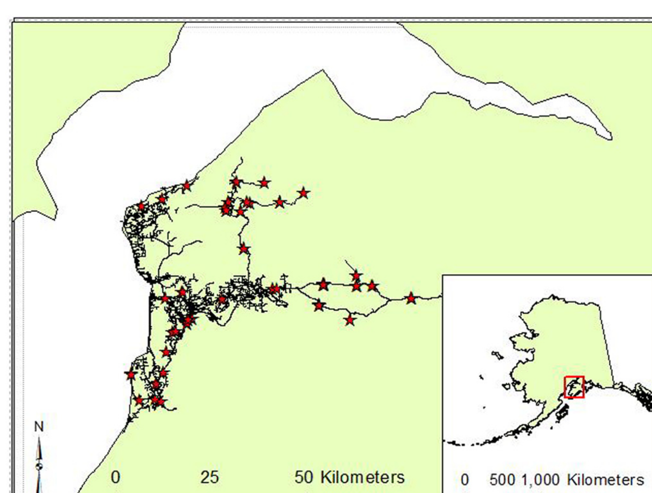


Fig. 1. Map of Alaska with study area (red box), Kenai Peninsula Borough Roads, and Study Sites where soil samples were collected (red stars).