

Navigating Data Linkages in Environmental Health Research with CHORDS: Connecting Health Outcomes Research and Data Systems

Aubrey K. Miller¹, Darius Bost^{1,3}, Adam Burkholder¹, Trisha Castranio¹, David Fargo¹, Ann Y. Liu^{1,2},
Kyle Messier¹, Alison Motsinger-Reif¹, David Reif¹, Charles Schmitt¹

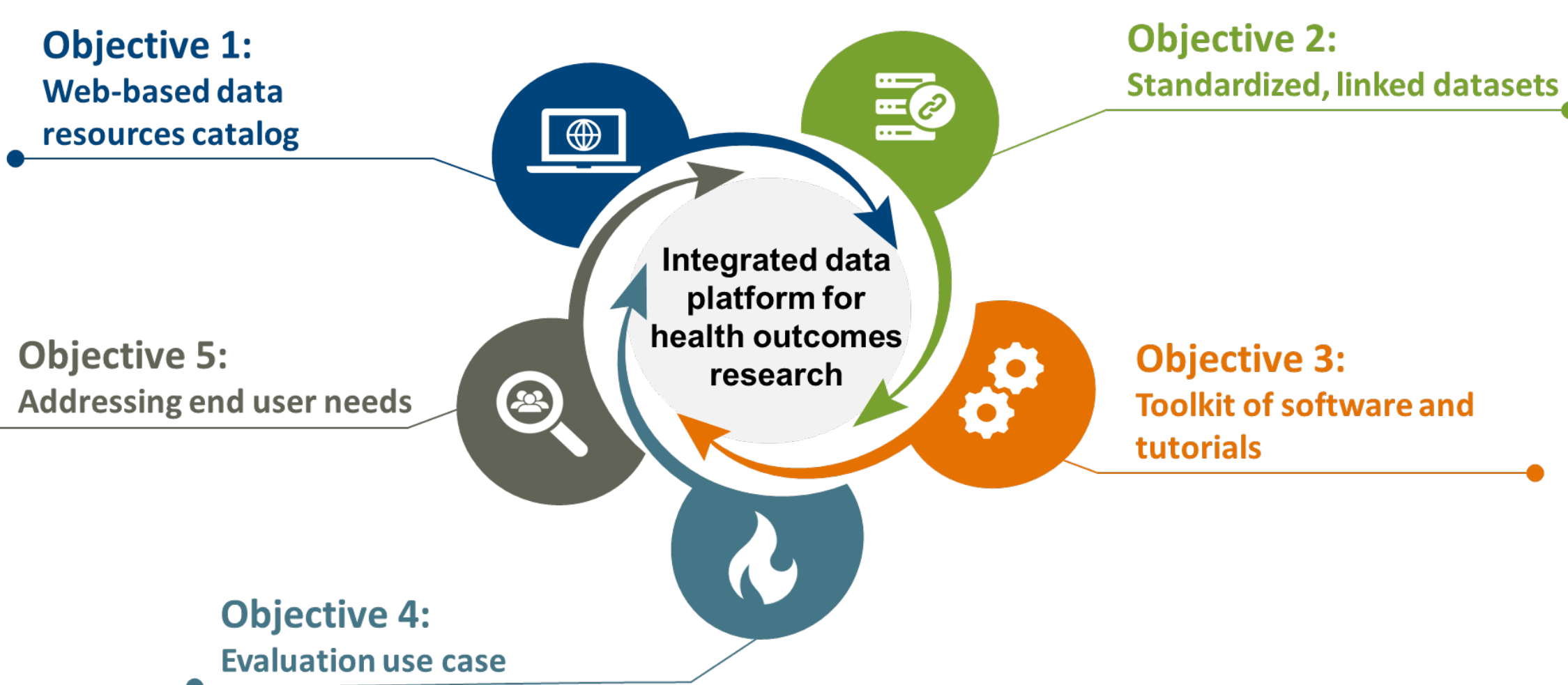
¹National Institute of Environmental Health Sciences, Research Triangle Park, NC;

²Kelly Government Solutions, Rockville; MD ³GAP Solutions, Durham, NC

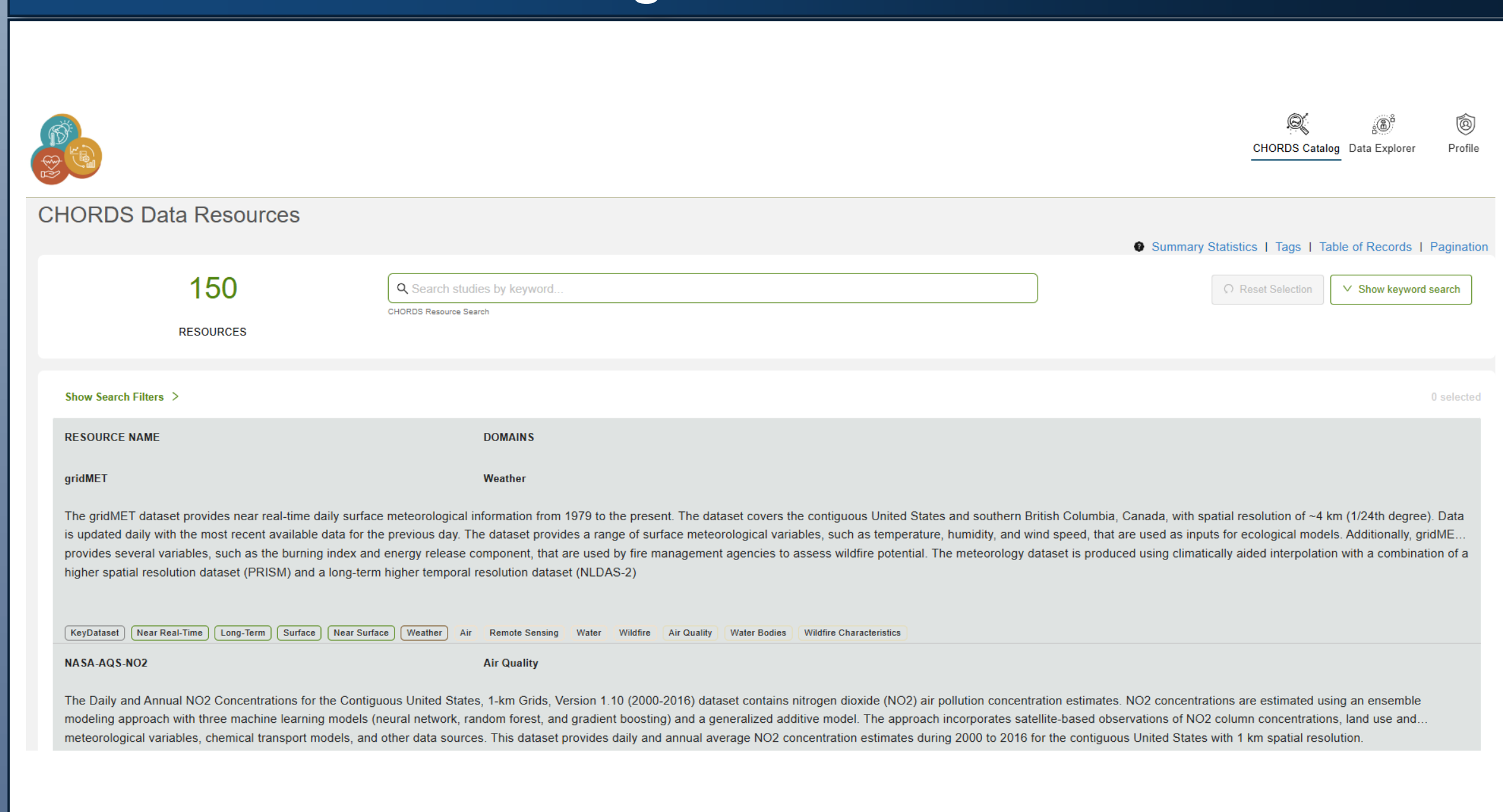
Background

The CHORDS project **strengthens data infrastructure to facilitate research connections** between environmental exposures and health outcomes. Exposures such as poor air and water quality, extreme temperatures, and natural disasters can cause both short-term health effects, like asthma attacks and heat stroke, and long-term conditions such as heart disease and cancer. By **providing accessible, interoperable data and analytical resources**, CHORDS supports researchers, health practitioners, and public officials in assessing environmental health risks, developing evidence-based interventions, and anticipating future health challenges to better protect communities.

CHORDS Objectives and Deliverables



Data Resources Catalog



CHORDS Data Resources

150 RESOURCES

Search datasets by keyword

Summary Statistics | Tags | Table of Records | Pagination

Show Search Filters >

0 selected

RESOURCE NAME

DOMAINS

gridMET Weather

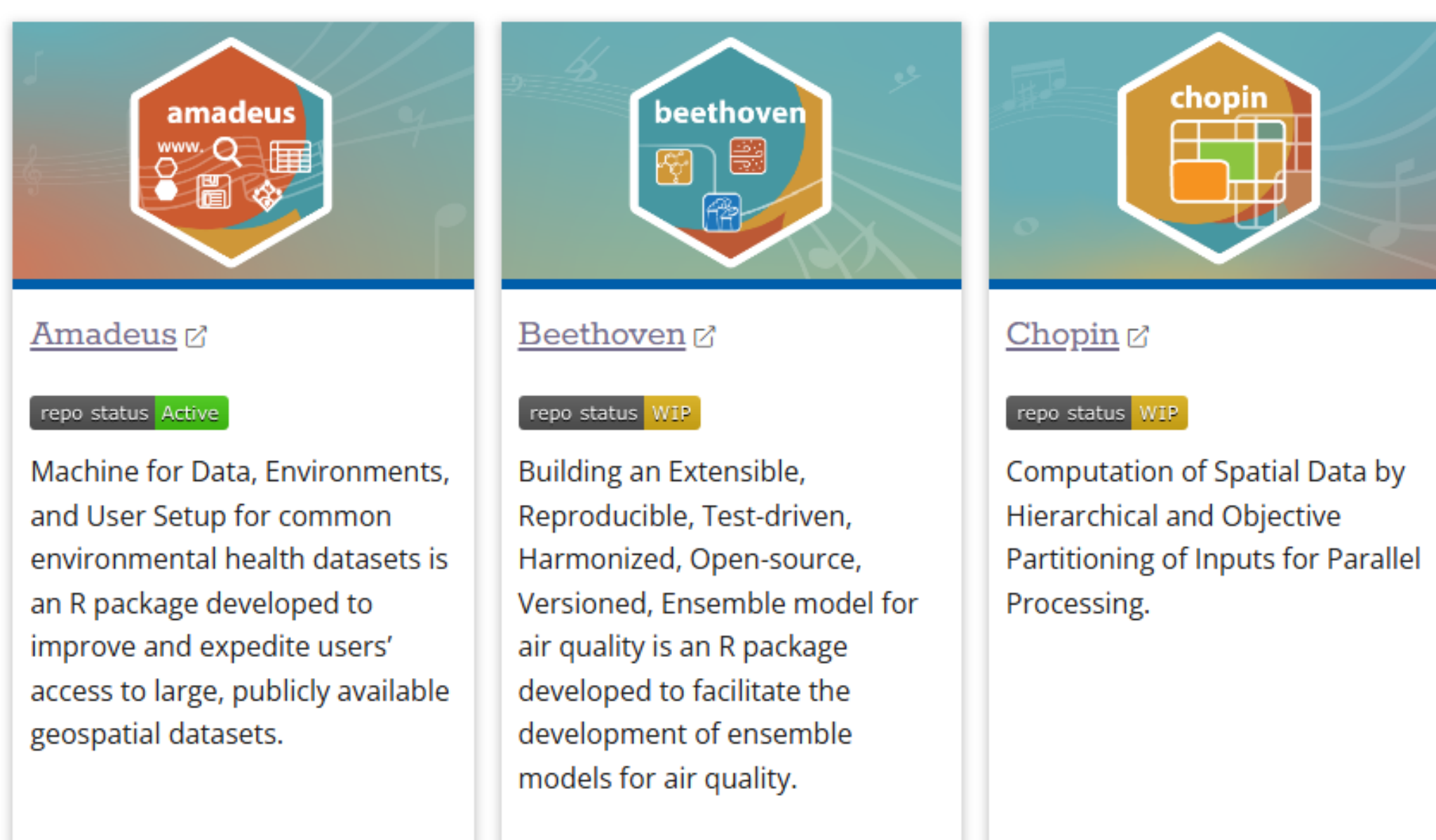
The gridMET dataset provides near real-time daily surface meteorological information from 1979 to the present. The dataset covers the contiguous United States and southern British Columbia, Canada, with spatial resolution of ~4 km (1/24th degree). Data is updated daily with the most recent available data for the previous day. The dataset provides a range of surface meteorological variables, such as temperature, humidity, and wind speed, that are used as inputs for ecological models. Additionally, gridMET... provides several variables, such as the burning index and energy release component, that are used by fire management agencies to assess wildfire potential. The meteorology dataset is produced using climatically aided interpolation with a combination of a higher spatial resolution dataset (PRISM) and a long-term higher temporal resolution dataset (NLDAS-2).

KeyDataset | Near Real-Time | Long-Term | Surface | Near Surface | Weather | Air | Remote Sensing | Water | Wildfire | Air Quality | Water Bodies | Wildfire Characteristics

NASA.AQS.NO2 Air Quality

The Daily and Annual NO2 Concentrations for the Contiguous United States, 1-km Grids, Version 1.10 (2000-2016) dataset contains nitrogen dioxide (NO2) air pollution concentration estimates. NO2 concentrations are estimated using an ensemble modeling approach with three machine learning models (neural network, random forest, and gradient boosting) and a generalized additive model. The approach incorporates satellite-based observations of NO2 column concentrations, land use and meteorological variables, chemical transport models, and other data sources. This dataset provides daily and annual average NO2 concentration estimates during 2000 to 2016 for the contiguous United States with 1 km spatial resolution.

Open-source Software Tools



Amadeus [www.amadeus.org](#)

repo status: Active

Machine for Data, Environments, and User Setup for common environmental health datasets is an R package developed to improve and expedite users' access to large, publicly available geospatial datasets.

Beethoven [www.beethoven.org](#)

repo status: WIP

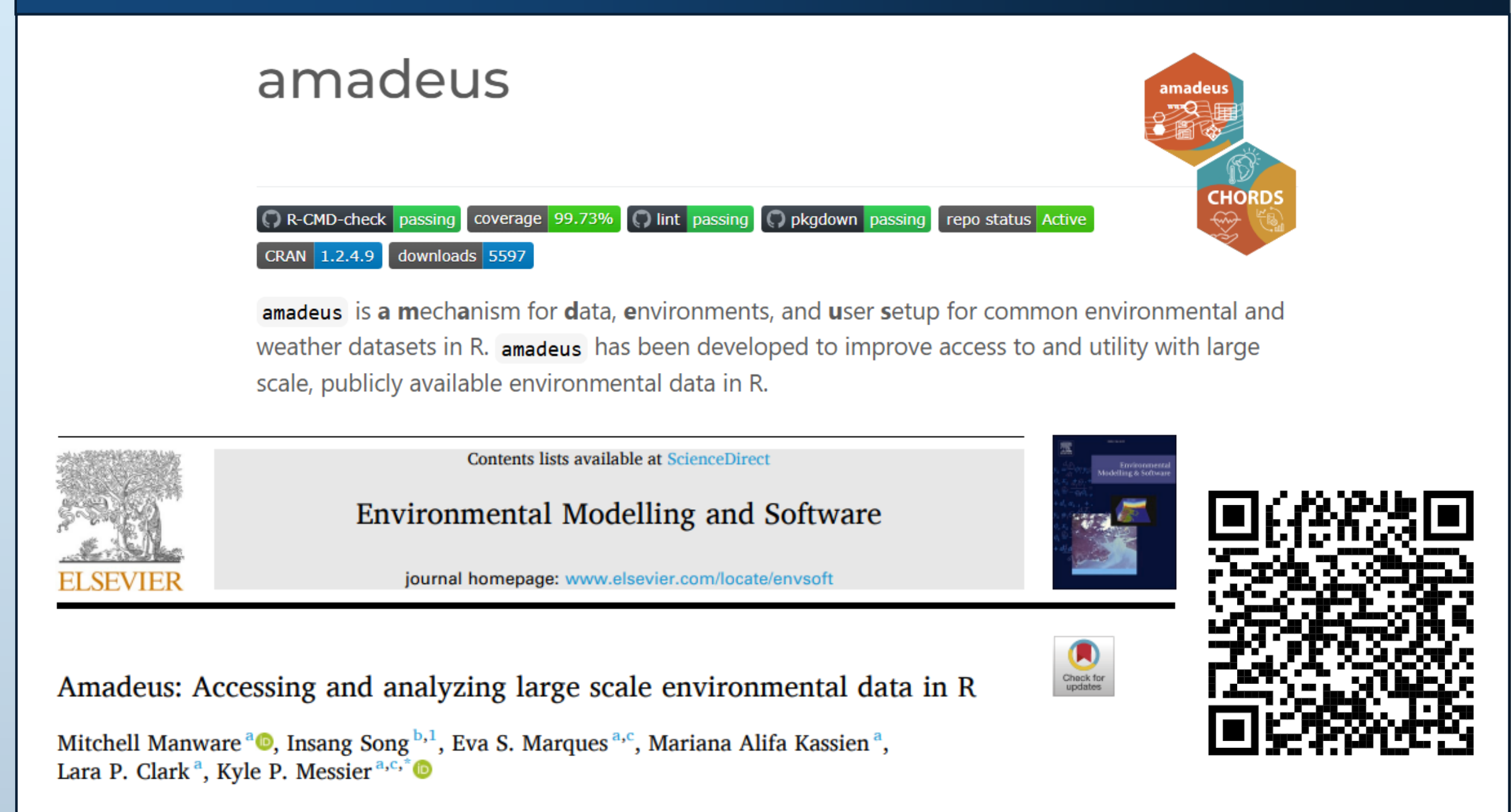
Building an Extensible, Reproducible, Test-driven, Harmonized, Open-source, Versioned, Ensemble model for air quality is an R package developed to facilitate the development of ensemble models for air quality.

Chopin [www.chopin.org](#)

repo status: WIP

Computation of Spatial Data by Hierarchical and Objective Partitioning of Inputs for Parallel Processing.

Adameus Software



amadeus

R-CMD-check: passing coverage: 99.73% lint: passing pkgdown: passing repo status: Active

CRAN 1.2.4.9 downloads 5597

amadeus is a mechanism for data, environments, and user setup for common environmental and weather datasets in R. amadeus has been developed to improve access to and utility with large scale, publicly available environmental data in R.

Contents lists available at ScienceDirect

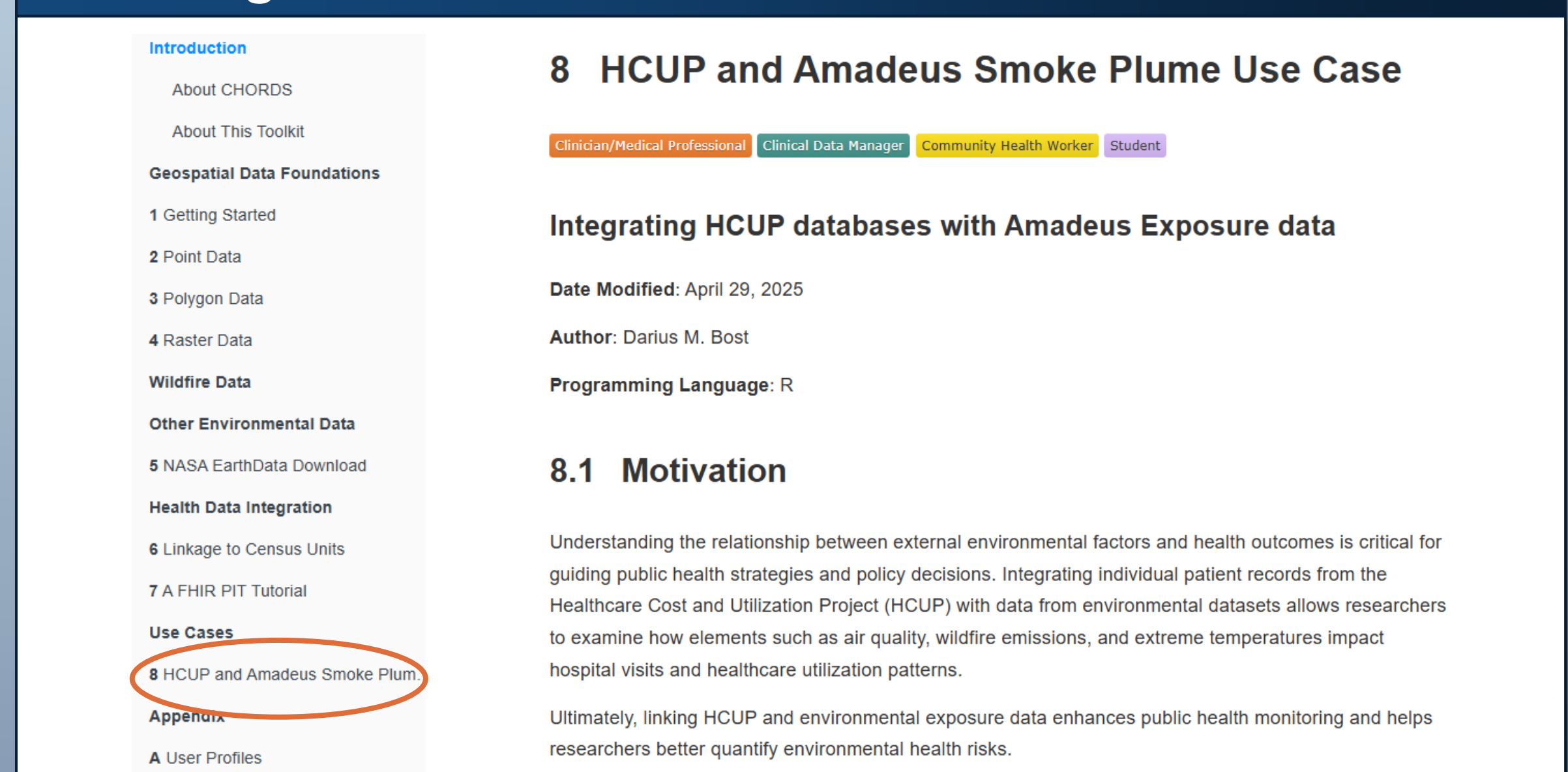
Environmental Modelling and Software

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

Amadeus: Accessing and analyzing large scale environmental data in R

Mitchell Manware^a, Insang Song^{b,1}, Eva S. Marques^{a,c}, Mariana Alifa Kassien^a, Lara P. Clark^a, Kyle P. Messier^{a,c}

Training and Tutorials



Introduction

About CHORDS

About This Toolkit

Geospatial Data Foundations

1 Getting Started

2 Point Data

3 Polygon Data

4 Raster Data

Wildfire Data

Other Environmental Data

5 NASA EarthData Download

Health Data Integration

6 Linkage to Census Units

7 A FHIR PIT Tutorial

Use Cases

8 HCUP and Amadeus Smoke Plume

Appendix

A User Profiles

8 HCUP and Amadeus Smoke Plume Use Case

Clinician/Medical Professional Clinical Data Manager Community Health Worker Student

Integrating HCUP databases with Amadeus Exposure data

Date Modified: April 29, 2025

Author: Darius M. Bost

Programming Language: R

8.1 Motivation

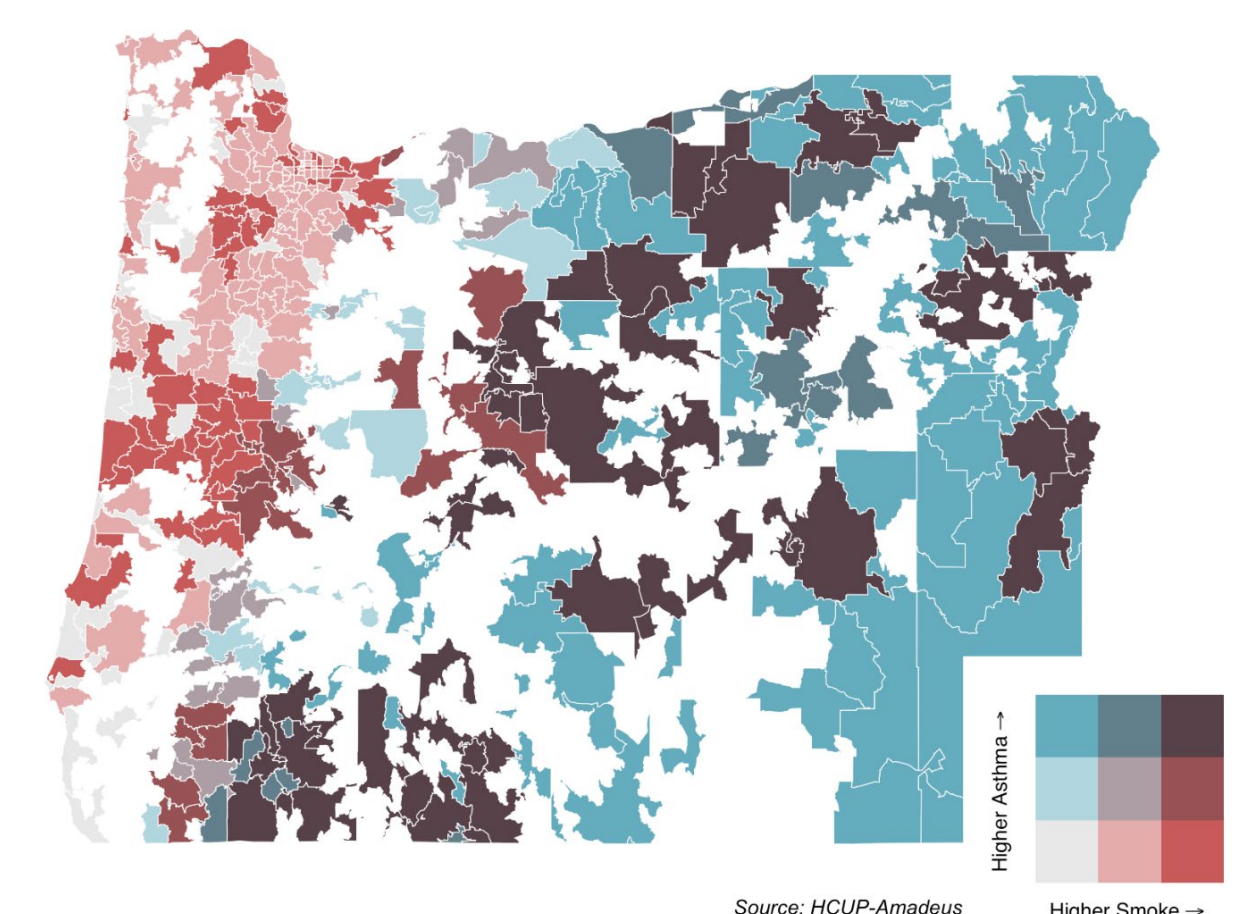
Understanding the relationship between external environmental factors and health outcomes is critical for guiding public health strategies and policy decisions. Integrating individual patient records from the Healthcare Cost and Utilization Project (HCUP) with data from environmental datasets allows researchers to examine how elements such as air quality, wildfire emissions, and extreme temperatures impact hospital visits and healthcare utilization patterns.

Ultimately, linking HCUP and environmental exposure data enhances public health monitoring and helps researchers better quantify environmental health risks.

Wildfire Use Case with HCUP Data (Oregon, 2021)

- Logistic regression model: exposure to medium and heavy smoke significantly increases the odds of asthma diagnosis
- Demonstrates utility of amadeus software in processing NOAA HMS wildfire smoke data
- Combining environmental and clinical data can support targeted public health interventions

Asthma Prevalence vs Heavy Smoke Exposure by ZIP Code
Bivariate map showing intersection of health and environmental burden



Conclusion

By building a data ecosystem of evolving publicly available resources, CHORDS aims to empower a diverse group of end users – including researchers, health care providers, policy and decisionmakers, and community groups – seeking to examine and mitigate the adverse health consequences of wildfires and other environmental health disasters and emergencies.



Supported by funding from the Office of the Secretary Patient-Centered Outcomes Trust Fund