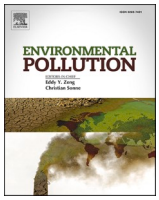
Environmental Pollution 342 (2024) 123105

Contents lists available at ScienceDirect

Environmental Pollution

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



# Understanding the driving mechanisms of site contamination in China through a data-driven approach☆

Kai Li [a,b], Ranhao Sun [a,b,*]

[a] State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China
[b] University of Chinese Academy of Sciences, Beijing, 100049, China

ARTICLE INFO

Handling Editor: Hefa Cheng

*Keywords:*
Site contamination
Driving mechanism
Key driver
Interpretable random forest model
Partial dependence plot

ABSTRACT

China currently faces significant environmental risks stemming from contaminated sites. The driving mechanism of site contamination, influenced by various drivers, remain obscured due to a dearth of quantitative methodologies and comprehensive data. Here, we used a data-driven causality inference approach to construct an interpretable random forest (RF) model. Results show that: (1) the trained RF model demonstrated remarkable predictive accuracy for identifying contaminated sites, with an accuracy rate of 0.89. In contrast to conventional correlation analysis, the RF model excels in discerning the key drivers through non-linear and genuine causal relationships between these drivers and site contamination. (2) Among the 25 potential drivers, we identified 18 key drivers of site contamination. These drivers encompass a broad spectrum of factors, including production and operational data, pollutant control level, site protection capability, pollutant characteristics, and physical-geographical conditions. (3) Each key driver exerts varying impacts on site pollution, with diverse directions, intensities, and underlying patterns. The partial dependence plots (PDPs) illuminate the role of each key driver, its critical value contributing to site pollution, and the interplay between these drivers. The key drivers facilitate the realization of three primary contamination processes: uncontrolled release, effective migration, and persistent accumulation. In light of our findings, environmental managers can proactively prevent site contamination by regulating single, dual, and multiple key drivers to disrupt critical pollution processes. This research offers valuable insights for devising targeted strategies and interventions aimed at mitigating environmental risks associated with contaminated sites in China.

## 1. Introduction

In recent years, there has been an increasing awareness of environmental protection and the severity of environmental problems in China (Wang et al., 2012), leading to significant attention being focused on the issue of contaminated sites (Liu et al., 2014; Yang et al., 2014). Contaminated sites not only disrupt soil ecosystems and contaminate groundwater (Balikova et al., 2022), but they also pose a substantial threat to human health and diminish the value of land (de Vor and de Groot, 2011; Fang et al., 2017; Zhang et al., 2022a,b). For instance, in 2015, a contaminated site associated with a chemical plant led to numerous students experiencing abnormal physical symptoms, including dermatitis and leukopenia. Despite China's efforts to upgrade industries and close polluting enterprises (Liu et al., 2017), there are still hundreds of thousands of operating enterprises facing the growing risk of site contamination (Zhang et al., 2022a,b). Consequently, the effective reduction of site contamination risk and the accurate implementation of measures to control environmental risks at these sites have become pressing concerns in China.

A comprehensive understanding of the primary processes and driving mechanisms behind site contamination is crucial for the successful implementation of prevention and control measures. Research has indicated that contaminated site occurrences are closely linked to improper production and handling of toxic and harmful substances, non-compliant emissions, leaks, and pollutant accumulations (Hou and Li, 2017; Zhao et al., 2018). Indicators reflecting enterprise management, such as pollutant emissions, waste management practices, and production technologies, significantly influence the development of

☆ This paper has been recommended for acceptance by Hefa Cheng.

* Corresponding author. State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China.
*E-mail address:* rhsun@rcees.ac.cn (R. Sun).

https://doi.org/10.1016/j.envpol.2023.123105
Received 16 August 2023; Received in revised form 24 November 2023; Accepted 4 December 2023
Available online 7 December 2023