

https://dspace.org/

Cropland associated with risk of Parkinsons disease in the northern Great Plains

B Krzyzanowski; K Beyene; S Searles Nielsen; J Killion; Brad Racette

https://hdl.handle.net/10539/45503
WOS

Downloaded from a DSpace repository. https://dspace.org/

Parkinsonism and Related Disorders 132 (2025) 107288

Contents lists available at ScienceDirect

Parkinsonism and Related Disorders

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

Cropland associated with risk of Parkinson's disease in the northern Great Plains

Brittany Krzyzanowski [a,*], Kassu M. Beyene [a], Susan Searles Nielsen [b], Jordan A. Killion [a], Brad A. Racette [a,b,c]

[a] Department of Neurology, Barrow Neurological Institute, 240 W Thomas Rd, Phoenix, AZ, 85013, USA
[b] Department of Neurology, Washington University School of Medicine, 660 S Euclid Ave, St. Louis, Missouri, 63110, USA
[c] School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, 27 St Andrews Rd, Parktown, 2193, South Africa

ARTICLE INFO

Keywords:
Parkinson
Epidemiology
Geographic
Environmental
Cropland density
PM2.5

ABSTRACT

Introduction: We sought to identify regional associations between cropland density and crop types and PD in the U.S.

Methods: We conducted a population-based study of 21,639,190 Medicare beneficiaries, 89,790 with incident PD in 2009. We used county-level geographic weighted regression (GWR) to identify region(s) of the U.S. where the association between PD RR and cropland density was strongest. In a broad region identified by GWR in which cropland density was associated with PD, we performed logistic regression using individual-level beneficiary data (2733 cases and 805,984 non-cases) with high-resolution cropland density data. We adjusted for age, sex, race, smoking, healthcare utilization, and $PM_{2.5}$ (particulate matter <2.5 μm). We then explored PD-cropland associations for *each type* of crop within a subregion, in which the association was the strongest.

Results: GWR identified a 9-state region in the Great Plains in which county-level cropland density and PD RR were associated. Within this region, the strongest GWR coefficients centered around the Williston Basin. High-resolution analysis demonstrated an association between cropland density within a 5-mile radius of residential zip+4 and PD. When comparing the highest to lowest quartile of cropland density, the odds ratio (OR) for PD was 1.14 (95 % confidence interval [CI] 1.01–1.27) in the 9-state region and 1.99 (95 % CI 1.09–3.61) in the Williston Basin. In the Williston Basin, percentage of sunflowers, winter wheat, and alfalfa within 5 miles of a beneficiary's zip+4 was associated with PD.

Conclusion: We identified a region-specific association between cropland and crop type and PD in the Williston Basin.

1. Introduction

The discovery of MPTP-induced dopaminergic neurotoxicity in humans and non-human primates fueled decades of research into pesticide-Parkinson's disease (PD) associations due to the structural similarities between MPTP and the herbicide paraquat [1]. Since pesticides vary in terms of their active ingredients and proposed mechanisms of neuropathogenesis, other pesticide neurotoxicants might contribute to PD risk, as well [2,3]. Epidemiologic studies have sought to investigate the relationship between occupational, environmental, and personal residential pesticide exposure and PD risk or progression [4–6], with occupational exposure being fairly consistently associated with PD. The association between environmental exposures and PD has yielded more mixed results likely due to methodological limitations, including use of self-reported exposure [4,5,7] or proxies for exposure, such as pounds of pesticide per acre [7], residential distance-to-farm [6], and rural/urban categorizations [8]. However, recent improvements to exposure assessment, including use of higher-resolution geographic data [7,9–11], allow epidemiological studies to better address the association between PD and environmental exposure to agricultural pesticides. Specifically, several new and powerful geographic information systems (GIS) and spatial analytical methods enable investigators to leverage large, high-resolution environmental datasets, including those from satellite imagery, remote sensing, and interpolated rasters. These data,

* Corresponding author. Department of Neurology, Barrow Neurological Institute, 240 W. Thomas Road, Phoenix, AZ, 85013, USA.
E-mail addresses: brittany.krzyzanowski@barrowneuro.org (B. Krzyzanowski), kassu.beyene@barrowneuro.org (K.M. Beyene), snielsen@wustl.edu (S.S. Nielsen), jordan.killion@barrowneuro.org (J.A. Killion), brad.racette@barrowneuro.org (B.A. Racette).

https://doi.org/10.1016/j.parkreldis.2025.107288
Received 25 September 2024; Received in revised form 13 December 2024; Accepted 13 January 2025
Available online 14 January 2025
1353-8020/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

as precise as 30 by 30-m pixels, can be used to estimate person-level environmental exposure to a feature of interest based on proximity to, or density around, an individual's home residence. Notably, high-resolution agricultural cropland data is available for the contiguous United States (U.S.) at spatial resolutions far more refined than used in most past environmental epidemiological studies of PD; although investigators have used high-resolution exposure data in geographically focused studies of agricultural pesticides in the Central Valley of California [7,9,11], this association in most of the U.S. remains unexplored. In this study, we use high-resolution land use data along with nationwide, population-based Medicare data to explore the relationship between cropland density and PD risk by utilizing advanced GIS and spatial methodology.

2. Methods

2.1. Standard protocol approvals, registrations, and patient consents

The Institutional Review Boards at Barrow Neurological Institute and Washington University in St. Louis determined that this study meets the 2018 Common Rule Exempt Category Secondary Research along with full waiver of HIPAA Authorization.

2.2. Study population

Our source population was all Medicare-aged United States citizens living in the contiguous United States. Our Medicare database population, contained 21,639,190 U S. beneficiaries from Medicare research files (CMS, 2009) who met initial study eligibility criteria as an incident PD case or comparable non-case (control) [12]. Incident PD cases included all eligible beneficiaries living with at least one International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis code of 332 or 332.0 in 2009, but no prior year. Controls did not have a code for PD. For both cases and controls, we used the following study eligibility criteria to ensure a population-based sample with complete data: age-eligible for Medicare at least two years before diagnosis/selection (age $\geq$66 years, 11 months), no Part C (Medicare Advantage/health maintenance organization) coverage, age $\leq$90, and U.S. residence, all in 2009, and without additional (e.g., medical) inclusion/exclusion criteria except Lewy body dementia and atypical parkinsonism [12]. Our Medicare database population included 89,790 incident PD cases and 21,549,400 controls from 2009 who met all above eligibility criteria. Our study population was based on those in the Medicare database population with geocodable address history two years prior to diagnosis (2007).

2.3. Study approach

2.3.1. Overview

We performed a population-based analysis of cropland and incident PD using a three-part approach. In our initial analysis we investigated cropland density as a proxy for pesticide and other agricultural exposures. This cropland density study first leveraged geographic methods to map the county-level relationship between cropland density and PD risk to identify a region for an individual-level case-control study. We then performed an individual-level case-control study using zip+4 data and high-resolution cropland density in a 5-mile radius (based on 100 $\times$ 100m cropland data) to formally test the relationships observed in our county-level geographic analysis. Finally, we used the individual case-control results to inform a third analysis in which we tested the association between PD risk and the different types of crops in a smaller geographic region which enabled the use of even higher resolution 30 $\times$ 30m cropland data.

2.3.2. County-level cropland density analysis

We performed geographic weighted regression (GWR) to obtain a map of linear regression beta coefficients from a model in which the dependent variable was a county-level spatially smoothed PD relative risk (RR) that adjusted for age, sex, race, smoking, and healthcare utilization and the independent variable was county-level cropland density (percent cropland), as described below. We derived the dependent variable (the spatially smoothed, adjusted RR) as detailed previously [12]. Briefly, we used the integrated nested Laplace approximation method for Bayesian inference in R-Project (R-INLA) version 4.1.2, and incorporated indirect age-sex-race-standardized incidence ratios for PD, spatial smoothing (using the conditional autoregressive distribution), and adjustment for healthcare utilization (the number of physician visits and outpatient visits per county) and current smoking (county-level prevalence of cigarette use) from the Institute for Health Metrics and Evaluation (IHME). We calculated county-level cropland density using zonal statistics in ArcMap and land use data from 2008 (the earliest available data) from the U.S. Department of Agriculture. We entered county-level cropland density into our GWR model as a single decile variable. Because there is no reference category, the GWR beta coefficients can be interpreted as the difference in PD RR when going from one decile of cropland density to the next. As a sensitivity analysis we entered cropland density into the model as a continuous variable ranging from 0 to 1, such that the GWR beta coefficients can be approximately interpreted as the difference in PD risk comparing complete (100 %) cropland density to no cropland exposure. We mapped the GWR coefficients and refined the map by computing 95 % confidence intervals (CIs), only retaining counties where both the CI excluded values of 0 and GWR coefficients were positive.

2.3.3. Individual-level cropland density case-control study

Based on the above geographic discovery methods, we performed a high-resolution, individual-level, case-control study for the region with the strongest cropland-PD association (largest significant beta coefficient). We used logistic regression with incident PD as the outcome and cropland density as the independent variable. We modeled cropland density as a continuous variable to obtain p-values for trend, and then in quartiles to allow for non-linear associations. Specifically, we defined cropland density as the proportion of cropland (based on 100 $\times$ 100m cropland data) within a 5-mile radius buffer around each beneficiary's zip+4 center. We used a 5-mile buffer based on previous research [13] and based on the level of precision of zip+4 in rural areas. In all models we adjusted *a priori* for beneficiary-level demographic information (age, sex, race), smoking probability [14], and measures of healthcare utilization in the year before diagnosis/reference—obtained from the Medicare BASF. We defined healthcare utilization as the number of inpatient days and outpatient visits for all beneficiaries in our study. We calculated predicted smoking probability for each beneficiary in our study using county-level prevalence of current smoking and individual-level data from the Medicare beneficiary annual summary file. These individual-level predictors included sex, race, birth cohort, and selected medical conditions (chronic obstructive pulmonary disease, lung cancer, stroke, acute myocardial infarction, other ischemic cardiovascular disease, stroke/transient ischemic attack, chronic kidney disease, osteoporosis, and depression). The gold standard variable on which we built this model replicates the well-established inverse association between smoking and PD and has been validated [12]. In addition to this basic model, we adjusted for other potential confounders in sensitivity analyses. We adjusted for average annual particulate matter with a diameter $<$2.5 μm ($PM_{2.5}$) from the Socioeconomic Data Application Center (SEDAC) from 1998 to 2000, a period of exposure before PD diagnosis. We extracted $PM_{2.5}$ values from the centroid of each zip+4 as previously [12], for use as a potential confounder in individual-level analyses. Because there is some evidence also supporting an association between urban living and PD [8], we also linked each beneficiary to their census tract level rural-urban commuting area (RUCA) code to adjust for differences between other neurotoxic exposures tied to urban or rural residence. For our adjustment variables, we modeled $PM_{2.5}$ as

two linear splines [12] and census tract level RUCA as a categorical variable with three categories. RUCA code 1 is considered urban, RUCA codes 2 through 6 are considered suburban, and codes 7 through 10 are rural. As a sensitivity analysis, we excluded major metropolitan urban populations (beneficiaries living in census tracts where RUCA code = 1 [metropolitan cores]) because these areas are generally not environmentally exposed to many pesticides that are only applied to crops. We tested (on the multiplicative scale) whether the cropland density-PD association differed by age in order to assess whether results from our Medicare-aged population could limit generalizability. As a sensitivity analysis, we adjusted for years living in the same zip code using the years of Medicare data available to us (2004–2009).

Volatilization of pesticides and wind erosion of contaminated soils may increase the distances pesticides travel [15]. Fine particles of 5 μm can drift up to three miles in very light wind conditions (3 mph) and even further with stronger winds [13]. For this reason, we also tested (on the multiplicative scale) for effect modification of the cropland-PD association by wind speed in logistic regression models. We used data on average annual windspeed at 10 m above surface level from 2007 to 2013 from the National Renewable Energy Laboratory. We hypothesized that any positive cropland-PD association would be more marked in areas with greater wind speeds. Therefore, we also used bivariate local indicators of spatial association (LISA) to assess and visualize local spatial correlation between the strength of the PD-cropland relationship and average annual windspeed. A bivariate LISA map delineates four cluster categories (high-high, low-low, low-high, and high-low). The high-high category represents counties where GWR coefficients for the PD-cropland relationship and windspeed are both high relative to their means, and the low-low category represents counties where GWR coefficients and windspeed are both low; these categories are consistent with a positive correlation between GWR coefficients and windspeed. The two discordant (low-high and high-low) categories are inconsistent with a positive correlation.

2.3.4. Crop type study

Within the same subregion identified from our geographic discovery methods, and same individual-level case-control study, we investigated the association between each of the different *types* of crops in that subregion and PD. For our crop type analysis, we used higher-resolution (30 × 30m) crop type data using the same source as for our overall cropland variable. This data source includes 103 potential crops throughout the U.S. Given the sparse geographic coverage of some of the 41 different crop types in our region of interest, we modeled the crop density in the 5 miles surrounding each beneficiary's Zip+4 center for each crop type as a continuous variable to avoid information loss through categorization. To investigate the relationship between incident PD and crop types, we employed multivariable binary logistic regression. Using a backward stepwise variable selection approach with the Akaike Information Criterion (AIC) as the selection criterion, we identified the most parsimonious model that explains the association between PD risk and these factors. This method iteratively removes the risk factors from the full model (which adjusted for age, sex, race, smoking, and use of healthcare utilization by forcing these variables into the model) to achieve the most parsimonious representation. Our final model adjusted for the same covariates as with our primary cropland analysis which included those listed above with the addition of average annual $PM_{2.5}$ and RUCA. To check the potential multicollinearity among the crop types, we calculated the variance inflation factor (VIF). As a sensitivity analysis, given the prevalence of crop rotations in the 9-state region, we performed our crop type analysis using cropland data from 2009. These statistical analyses were performed on R (version 4.4.1; R Foundation for Statistical Computing).

3. Results

3.1. Nationwide cropland density and PD

For county-level mapping, our study population included 89,390 (99 %) of PD cases and 21,086,990 (98 %) of controls with county level information from our Medicare database population. Of those, we had 65,180 (73 %) cases and 15,561,435 (72 %) non-cases who could be successfully geocoded to the zip+4 level for use in our high-resolution individual-level case control study (Fig. S3). We observed a greater percentage of beneficiaries without geocodable zip+4 location data in the south and west compared to the north and east. The nationwide geospatial variation in the pattern of zip+4 information loss was generally similar by case status. (Fig. S4). In our GWR analysis we identified a significant positive association between county-level cropland density (Fig. S1) and county-PD RR that spanned much of the Great Plains Region in the U.S. (Fig. 1a). The positive GWR coefficients stretched across nine states (Idaho, Montana, North Dakota, South Dakota, Wyoming, Nebraska, Utah, Colorado, and Minnesota), with the strongest coefficients centering on the border between Montana and North Dakota in close overlap with a region called the Williston Basin and extending into the Powder River Basin of Wyoming (Fig. 1a). Modeling cropland density continuously confirmed a strong relationship between cropland density and PD risk in the northern Great Plains region (Fig. S2), including a majority of the Williston Basin and the Powder River Basin (Fig. 2).

3.2. High-resolution case-control study of cropland density in the 9-state region and the Williston Basin

Our nationwide cropland density-PD association study informed a high-resolution case-control study that included the nine states that contained the highest GWR coefficients. In this 9-state region, we identified 2733 incident PD cases and 805,984 controls, and Table 1 shows the demographic distribution of these beneficiaries. From the table, about 44 (55) percent of cases (controls) are female beneficiaries, and around the same fraction (96 %) of beneficiaries in the case and control groups are white. In addition, the age distribution of beneficiaries in case and control groups is very similar [case: mean (SD) = 78 (6.1), and control: mean (SD) is 76(6.2)]. When comparing the highest to the lowest quartile of cropland density, the odds ratio (OR) for PD was 1.14 (95 % CI 1.02–1.26) (Table 2). After restricting to non-metropolitan populations (1324 incident PD cases and 377,287 controls), the OR for our 9-state region analysis increased to 1.46 (95 % CI 1.18–1.81) (Table S1). Overall results were unchanged with adjustment for $PM_{2.5}$ or RUCA code. When we further restricted our analysis to the counties within the Williston Basin, the OR was 1.99 (95 % CI 1.09–3.61). Although not statistically significant, we found a trend, wherein the $PM_{2.5}$-PD association differed by age ($p_{interaction} = 0.09$), specifically such that the association was stronger in younger beneficiaries (results not in shown). Adjusting for years living in the same zip code did not change in our results (Table S2).

However, we did not detect effect modification by windspeed ($p_{interaction} = 0.42$) (results not shown). Contrary to these results, our bivariate LISA revealed local spatial correlation characterized by high-high clusters (counties where the strength of the PD-cropland density relationship was above average and wind speed was also above average) in the northern Great Plains Region and low-low clusters (counties where both the strength of the PD-cropland density relationship and windspeed were below average) in some South and Atlantic states (Fig. 2). Only a small number of discordant areas were observed.

3.3. Crop Type-PD associations in the Williston Basin

Of the 41 different crop types mapped in the Williston Basin in 2008, using a backward stepwise variable selection approach, six crop types

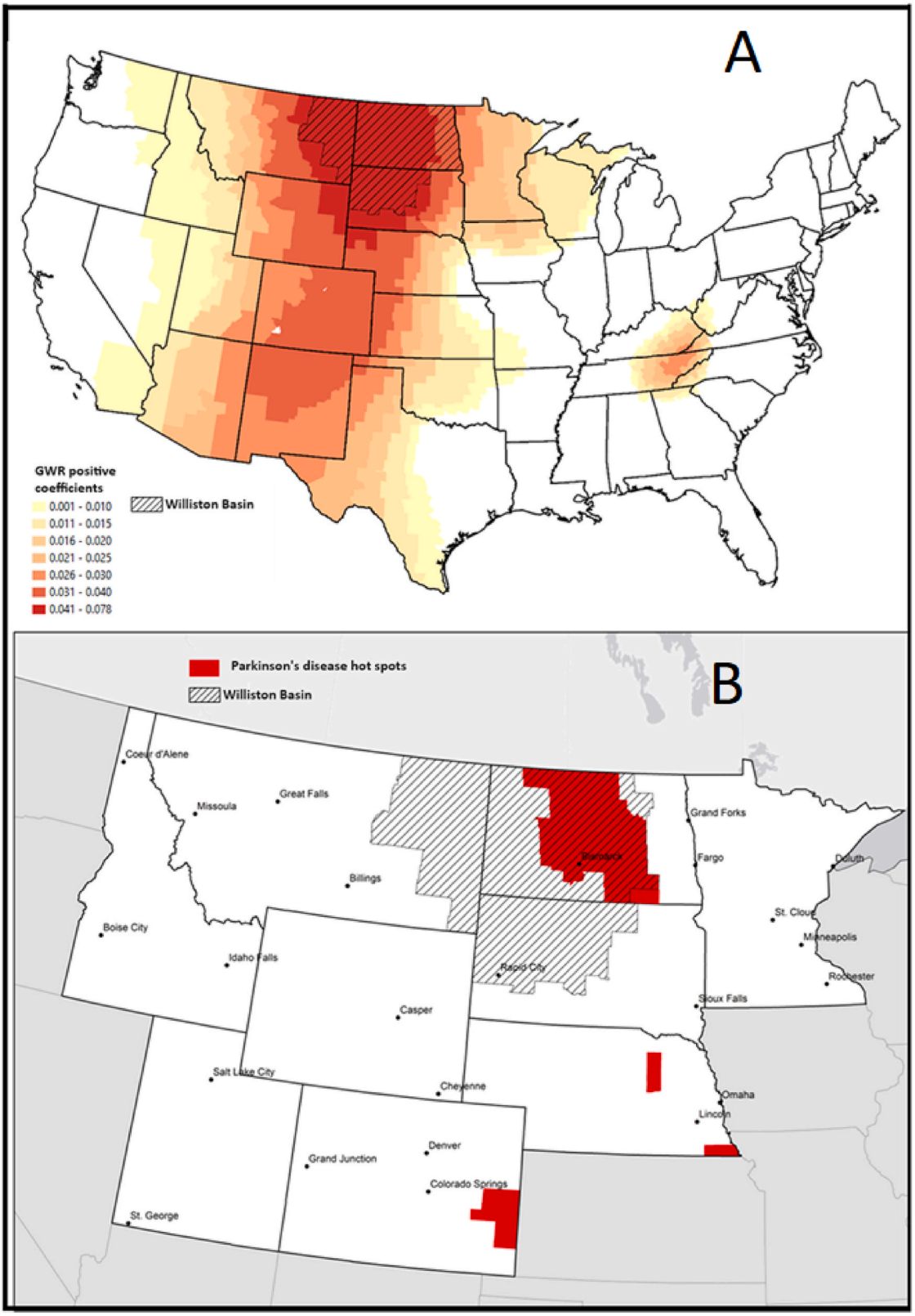


Fig. 1. GWR coefficient map showing location and strength of significant positive coefficients for the association between county-level cropland density and Parkinson disease relative risk in the U.S. that account for age, sex, race, smoking, healthcare utilization, and spatial dependency. Each coefficient represents the absolute difference in the relative risk with each additional decile of percent cropland exposure (Fig. 1a). The most marked associations appear in the Williston Basin, which contains a Parkinson disease hotspot [12] (Fig. 1b) and relatively high county-level cropland density compared to the rest of the nation.
GWR = geographically weighted regression; U.S. = United States.

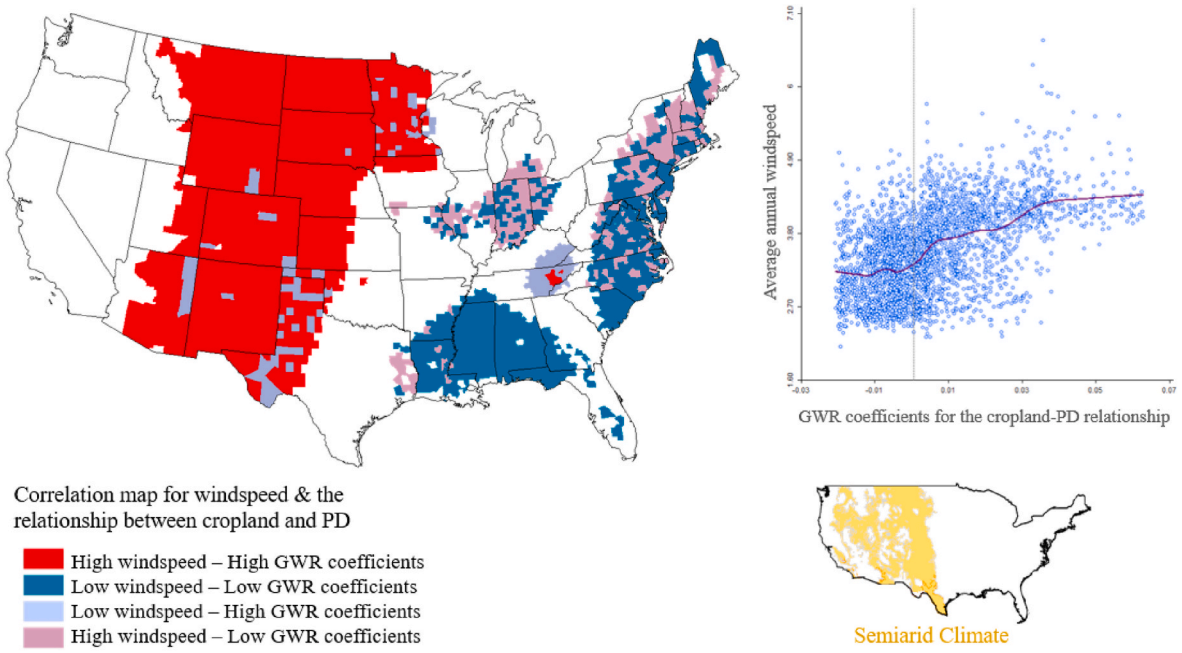


Fig. 2. Cluster map of the PD-cropland relationship among U.S. Medicare beneficiaries in 2009 overlaid with hot and cold spots of average annual windspeed from 2007 to 2013 indicates high concordance between the magnitude of positive cropland-PD associations and windspeed. Clusters were identified with bivariate LISA using: 1) geographic weighted regression coefficients for the relationship between county-level PD risk and cropland density that adjust for age, sex, race, smoking, healthcare utilization, and spatial dependency; and 2) average annual wind speed. The high-high category represents counties where the strength of the PD-cropland density relationship is above average and wind speed is also above average. The low-high and high-low clustering categories describe spatial outliers where a low-high county is one where the strengthen of the PD-cropland relationship is below average and windspeed is above average. The low-low category represents counties where the strength of the PD-cropland relationship and windspeed are both lower than average. The high-high clustering overlaps with the semiarid climate region of the U.S. where wind erosion and evaporation make pesticide droplets more susceptible to wind forces [25–27]. Locally weighted scatterplot smoothing curve for the relationship between average annual windspeed and the strength of the cropland-PD relationship confirms a positive relationship between the cropland-PD association and windspeed. Each dot represents a county (N = 3107).
PD = Parkinson's disease; LISA = local indicators of spatial association; U.S. = United States.

Table 1
Characteristics of incident PD cases and controls in the 9-state region, U.S. Medicare 2009.

	PD cases N = 2733	Controls N = 805,984
Female, %	44 %	55 %
Race/ethnicity, %		
White	96 %	97 %
Black	1 %	1 %
Hispanic	0.8 %	0.8 %
Asian	0.5 %	0.7 %
Native American	0.6 %	0.4 %
Pacific Islander/other/unknown	0.5 %	0.4 %
Age, mean (standard deviation)	78 (6.1)	76 (6.2)
Smoking probability (index), mean (SD)[a]	1.4 (0.5)	1.5 (0.5)
Average annual $PM_{2.5}$ ($\mu g/m^3$), mean (SD)	8.5 (1.7)	8.5 (1.8)
Health care utilization (average number of outpatient visits and inpatient days), mean (SD)	13.8 (12.1)	9.5 (10.2)

Nine-state region includes Idaho, Montana, North Dakota, South Dakota, Wyoming, Nebraska, Utah, Colorado, and Minnesota.
PD = Parkinson's disease; SD = standard deviation; U.S. = United States.
[a] Median predicted probability of ever vs never smoking divided by the number of diagnosis codes. Predicted probability was based on sex, race/ethnicity, birth cohort, and 661 diagnosis and procedure codes including codes specific to tobacco, codes for heavy alcohol use, codes for cancers with high attributable risk because of smoking (lung and larynx), and codes related to other conditions that were associated with smoking in the 2009 Behavioral Risk Factor Surveillance System (BRFSS) data among Americans aged 66–90 years.

were entered into our final fitted models (with and without adjusting for average annual $PM_{2.5}$ and RUCA, as results were similar) (Table 2). We assessed for multicollinearity among predictor variables in the final model and found no evidence of a significant problem, as the calculated VIF of all variables is less than 2 (results not shown). These results revealed a positive relationship between risk of PD and the density of three different crops in the five miles surrounding a beneficiary's zip+4 center: sunflowers, winter wheat, and alfalfa. For each of these three crops, the OR was relatively similar to the PD OR for cropland in general, with an approximately 20 % increased risk estimated for each unit increase in density of these crops within five miles of an individual's zip+4 (Table 2). As a sensitivity analysis, we performed our crop type analysis using a different year of crop data and found that results remained consistent (results not shown).

4. Discussion

In this nationwide, population-based geographic study, we identified a subregion in the Great Plains Region of the U.S. with a marked dose-response relationship between county-level cropland density and PD. In this 9-state region, we confirmed that cropland density within five miles of each beneficiary's residential zip+4 also was associated with PD risk with an estimated 24 % greater risk in a hypothetical comparison between those with 100 % cropland density to no cropland. Notably, within the 9-state region, we found the strongest association between cropland density and PD in the Williston Basin, which has relatively high risk of PD (in central North Dakota specifically) (Fig. 1b) (Fig. S5) [12] and relatively high county-level cropland density compared to the rest of the nation. In the Williston Basin, the risk of PD was nearly two-fold greater for those living in residences proximate to the greatest cropland density in the five miles surrounding their zip+4 location compared to those living in residences with the lowest cropland density. Our crop type investigation suggests that pesticides applied to sunflowers, winter

Table 2
Mean cropland density within 5-mile radius of zip+4 and risk of PD in the 9-state region, overall and restricted to the Williston Basin and crop type within 5-mile radius of zip+4 and risk of PD in the Williston Basin.

		Cases N = 2733	Controls N = 805,984	Basic Model OR (95 % CI)[a]	Also adjusted for $PM_{2.5}$ & RUCA code OR (95 % CI)[a,b,c]
9-State Region Cropland					
Continuous[d]		2733	805,984	1.24 (1.06, 1.45)	1.24 (1.03, 1.48)
p_{value}				0.008[e]	0.002[e]
Cropland Density Quartile, median (range)					
Q1	0 % (0–1%)	629	200,882	Reference	Reference
Q2	5 % (2–11 %)	649	199,505	1.02 (0.92, 1.14)	1.03 (0.92, 1.15)
Q3	22 % (12–37 %)	676	204,002	1.03 (0.92, 1.14)	1.03 (0.92, 1.15)
Q4	54 % (38–89 %)	779	201,595	1.14 (1.02, 1.26)[e]	1.14 (1.01, 1.28)[e]
p_{trend}				0.0001[f]	0.0001[f]
Williston Basin Cropland					
Continuous[d]		157	39,948	2.80 (1.31, 5.97)	2.76 (1.03, 7.39)
p_{value}				0.008[e]	0.043[e]
Cropland Density Quartile, median (range)					
Q1	4 % (0–9%)	20	8867	Reference	Reference
Q2	12 % (10–17 %)	41	10,971	1.50 (0.88, 2.56)	1.50 (0.87, 2.56)
Q3	34 % (18–40 %)	42	10,023	1.58 (0.92, 2.70)	1.53 (0.85, 2.75)
Q4	51 % (41–86 %)	54	10,107	2.06 (1.23, 3.46)[e]	1.99 (1.09, 3.61)[e]
p_{trend}				0.0001[f]	0.0001[f]
Williston Basin Crop Type[g,h]				**Full Model**	**Final Model**
Continuous[d]	**Sunflower**			**1.248 (1.065, 1.46)**[e]	**1.243 (1.051, 1.47)**[e]
	Winter wheat			**1.170 (1.041, 1.32)**[e]	**1.169 (1.037, 1.32)**[e]
	Millet			0.701 (0.456, 1.08)	0.698 (0.454, 1.07)
	Safflower			0.559 (0.175, 1.78)	0.561 (0.172, 1.83)
	Alfalfa			**1.215 (1.045, 1.41)**[e]	**1.211 (1.028, 1.43)**[e]
	Camelina			0.737 (0.376, 1.45)	0.737 (0.374, 1.45)

$PM_{2.5}$ = particulate matter <2.5 μm in diameter; PD = Parkinson's disease; μg/m^3 = micrograms per cubic meter; CI = confidence interval; OR = odds ratio; RUCA = rural urban commuting area.

[a] ORs and associated 95 % CIs from logistic regression with PD as the outcome and cropland density as the independent variable, adjusted for individual-level variables (age, sex, race, probability of ever/never smoking, use of healthcare utilization), based on cases and controls with geocodable zip+4 data for residence two years prior to diagnosis or reference.

[b] ORs also adjusted for average annual $PM_{2.5}$ and RUCA.

[c] Cases and controls were assigned the RUCA value of the census tract that corresponded to their zip+4 center. Excludes 289 beneficiaries with unknown RUCA.

[d] OR is per percent of land that is cropland within the 5-mile radius of zip+4 center.

[e] $p < 0.05$.

[f] The p-value for trend using a postestimation command for contrast for the basic model and fully adjusted model.

[g] Missing 1764 beneficiaries without smoking and health care utilization information.

[h] Using multivariable binary logistic regression with a backward stepwise variable selection approach.

wheat, and alfalfa (and/or their rotation crops) may be of interest for further examination as risk factors in that region.

There are several unique aspects of our study that distinguish our work from prior studies. We leveraged powerful GIS methods to perform computationally intensive spatial analyses on high-resolution data on cropland in a large, multistate region. Previous studies of agricultural pesticide exposure and PD have been limited to smaller geographic regions and populations [2,11,16]. We used a large, nationwide dataset to identify a region to focus on within our case-control with nearly 3000 incident PD cases. Furthermore, we included a sensitivity analysis that adjusted for $PM_{2.5}$ as a confounding factor because we have previously shown that $PM_{2.5}$ is related to PD [12]. Since adjustment did not affect results, $PM_{2.5}$ is unlikely to mediate the association between cropland and PD. Moreover, given that our Bivariate LISA map revealed that the cropland-PD association was stronger in regions that were windier, these additional results suggest that inhaled airborne pesticides might be relevant to PD risk, not just ingested pesticides from drinking water.

Our study adds to a growing literature implicating pesticide exposure as a risk factor for PD. Some of strongest evidence linking environmental pesticide exposure to PD comes from studies in California's Central Valley [2,7,10,11,16]. These studies utilize pesticide-use reporting (PUR) data available for the State of California, which provides information on pesticide application at the township level (6 × 6 mile). Additionally, many of the Central Valley studies augment their PUR data by integrating land use data on the exact location of crop fields to refine exposure estimates to smaller geographic regions (1–6 square miles around residential and workplace addresses) [7,9–11]. This augmented, high-resolution PUR data have been validated using metabolomics in a cohort of 176 older adults [17], however, it is only available for California. In contrast, the cropland density variable is available for the entire contiguous U.S. Like the prior work in California, we used the exact location of crop fields, however, our primary approach, which utilized cropland as our exposure variable, does not integrate information on active ingredient, probability, and frequency of pesticide application. Nonetheless, cropland density might be a reasonable proxy for exposure to specific pesticides in regions with homogenous crop types and farming practices, such as large portions of the Midwest and Great Plains. We found the strongest association between cropland density and PD in the Williston Basin which includes the North Dakota-Montana border, a region that has relatively high cropland density and relatively low cropland diversity [18,19]. The relatively low cropland diversity in this region may have enhanced our ability to detect an association when using cropland density, insofar as crop homogeneity leads to homogeneity in the specific agricultural pesticides used. In contrast, heterogeneity of crop types in certain parts of the nation might explain why we did not identify a strong association between cropland density and PD in some areas where previous studies have found associations between pesticide exposure and PD [2,11,16]. For instance, the Central Valley of California is located within a region known to have the highest diversity of crop types in the nation [18,19] and therefore our cropland density measure—if acting as a proxy for specific pesticides—would be less precise insofar as these specific pesticides are not used on all crop types.

In particular, we found that living near fields in which sunflowers, winter wheat, or alfalfa were planted was associated with increased risk of PD in the Williston Basin. Of interest, chlorpyrifos, an organophosphorus insecticide previously linked to PD [2], is commonly applied to these sunflowers, winter wheat, and alfalfa crops in the Williston Basin

[20]. Chlorpyrifos has been linked to disrupted dopaminergic neurotransmission and neuronal damage in the substantia nigra in animal models [3] and increased risk of PD in epidemiological studies [2]. That said, in addition to chlorpyrifos, dozens of other pesticides are commonly applied to sunflowers, winter wheat, and alfalfa in the Great Plains [21–24]. Of those pesticides, paraquat [21], permethrin [22], 2, 4-D [23], and atrazine [24] have been previously linked to PD. Nevertheless, the data on crop types are available at a higher resolution than pesticide data so future studies in this region will be needed to investigate the role of specific pesticides.

The semiarid climate of the Great Plains region promotes wind erosion and evaporation of aerosolized liquid particles, thereby reducing pesticide droplet size and making them more susceptible to wind forces [25–27]. Therefore, we speculate that the potential for pesticide drift (and the distance of drift) may be increased in the northern Great Plains region. Additionally, the northern Great Plains region has engaged in no-till dryland cropping with increased frequency since 1989 [28]. No-till dryland cropping is a no-irrigation farming practice that preserves soil moisture by conserving tillage [28] and thus often relies more heavily on pesticides compared to conventional farming which removes weeds during tilling. In the Williston Basin specifically, the U.S. Department of Agriculture reported that no-till small grain production has "increased from less than 5 % of the total cultivated land area in 1989 to nearly 25 % in 2004" [28]. We speculate that the combination of the hot, dry climate, higher than average windspeeds, and no-till agricultural practices may increase the potential for pesticide exposure in the northern Great Plains region. That said, northeastern Wyoming and southeastern Montanna are also a part of the "Powder River Basin" which is known for its coal mining reserves. Therefore, it's possible that air pollution (primarily sulfur dioxide) from power plant emissions may have contributed to the risk of PD in this area. Nevertheless, studies of the role of sulfur dioxide and PD risk have yielded mixed results [29,30].

We acknowledge some limitations to our study. First, cropland density is a proxy for amount of pesticide exposure, which we speculate might contribute to the risk of PD in this region. The use of this geographical surrogate for individual level exposure to pesticides introduces exposure measurement error, which likely is similar to cases and controls and therefore anticipated to likely have biased the reported associations toward the null. Despite using sophisticated geospatial statistics, there is spatial uncertainty when using zip+4 locations as proxies for home residences in rural areas since the true residence can sometimes be miles from the zip+4 center, further enhancing exposure measurement error and associated attenuation of results. That said our analysis that excluded metropolitan cores provided stronger results thus supporting the relationship between cropland density and PD risk. We were unable to assess the impact of duration of exposure as we were limited to cropland density based on the earliest available satellite imagery in 2008, which does not allow us to capture the long PD prodromal period, however we anticipate that cropland density in 2008 would be reasonably correlated with cropland density in previous years especially in this rural region which is relatively homogenous in terms of crop type. Additionally, because our crop type analysis was based on one year of data, it is possible that the pesticides applied to the rotation crops for sunflowers, winter wheat, and alfalfa could explain the association we observed; however, our sensitivity analysis revealed similar results when using a different year of cropland data. Another limitation, which applies to population-based Medicare data broadly, is the restriction on coverage to patients aged >65, and we observed possible effect modification of the cropland-PD association by age, however the association was stronger in younger beneficiaries. Finally, we also note that the validity of our data on PD diagnosis relies on competent diagnoses and data entry by healthcare providers and staff. Despite limitations, our study advances current knowledge of agricultural crops and different crop types in relation to PD risk and demonstrates the utility of using a multimethod geographic approach for investigating environmental risk factors.

In summary, by using state-of-the-art geographic techniques and high-resolution exposure measures, we identified a region-specific association between cropland density and PD in the U.S. Great Plains Region, possibly related to pesticides commonly applied to sunflowers, winter wheat, and alfalfa. Future work should investigate these specific pesticides and others used in that region.

CRediT authorship contribution statement

Brittany Krzyzanowski: Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Kassu M. Beyene:** Writing – review & editing, Methodology, Formal analysis. **Susan Searles Nielsen:** Writing – review & editing, Methodology. **Jordan A. Killion:** Writing – review & editing. **Brad A. Racette:** Writing – review & editing, Project administration, Funding acquisition.

Data availability

CMS does not permit data sharing under the data use agreement.

Funding

This work was supported by The Michael J. Fox Foundation for Parkinson's Research [grant number MJFF-000939]; the U.S. Department of Defense [grant number PD190057]; and the National Institutes of Health – National Institute of Environmental Health Sciences [grant number K01ES028295]. None played a role in the study design, collection, analysis, interpretation of data, writing of the report, or the decision to submit this article for publication.

Declaration of competing interest

This work was supported by The Michael J. Fox Foundation for Parkinson's Research [grant number MJFF-000939]; the U.S. Department of Defense [grant number PD190057]; and the National Institutes of Health – National Institute of Environmental Health Sciences [grant number K01ES028295]. Additional support was received from the Kemper and Ethel Marley Foundation, the Barrow Neurological Foundation, and the Moreno Family. None played a role in the study design, collection, analysis, interpretation of data, writing of the report, or the decision to submit this article for publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.parkreldis.2025.107288.

References

[1] J.W. Langston, The etiology of Parkinson's disease with emphasis on the MPTP story, Neurology 47 (1996) S153–S160, https://doi.org/10.1212/wnl.47.6_suppl_3.153s.
[2] N.M. Gatto, M. Cockburn, J. Bronstein, A.D. Manthripragada, B. Ritz, Well-water consumption and Parkinson's disease in rural California, Environ. Health Perspect. 117 (2009) 1912–1918, https://doi.org/10.1289/ehp.0900852.
[3] J. Zhang, H. Dai, Y. Deng, J. Tian, C. Zhang, Z. Hu, G. Bing, L. Zhao, Neonatal chlorpyrifos exposure induces loss of dopaminergic neurons in young adult rats, Toxicology 336 (2015) 17–25, https://doi.org/10.1016/j.tox.2015.07.014.
[4] J.A. Firestone, T. Smith-Weller, G. Franklin, P. Swanson, W.T. Longstreth Jr., H. Checkoway, Pesticides and risk of Parkinson disease: a population-based case-control study, Arch. Neurol. 62 (2005) 91–95, https://doi.org/10.1001/archneur.62.1.91.
[5] C.M. Tanner, G.W. Ross, S.A. Jewell, R.A. Hauser, J. Jankovic, S.A. Factor, S. Bressman, A. Deligtisch, C. Marras, K.E. Lyons, G.S. Bhudhikanok, D.F. Roucoux, C. Meng, R.D. Abbott, J.W. Langston, Occupation and risk of parkinsonism: a multicenter case-control study, Arch. Neurol. 66 (2009) 1106–1113, https://doi.org/10.1001/archneurol.2009.195.
[6] M.E. Hugh-Jones, R.H. Peele, V.L. Wilson, Parkinson's disease in Louisiana, 1999-2012: based on hospital primary discharge diagnoses, incidence, and risk in

relation to local agricultural crops, pesticides, and aquifer recharge, Int. J. Environ. Res. Publ. Health 17 (2020), https://doi.org/10.3390/ijerph17051584.
[7] A. Wang, S. Costello, M. Cockburn, X. Zhang, J. Bronstein, B. Ritz, Parkinson's disease risk from ambient exposure to pesticides, Eur. J. Epidemiol. 26 (2011) 547–555, https://doi.org/10.1007/s10654-011-9574-5.
[8] W. Koller, B. Vetere-Overfield, C. Gray, C. Alexander, T. Chin, J. Dolezal, R. Hassanein, C. Tanner, Environmental risk factors in Parkinson's disease, Neurology 40 (1990) 1218–1221, https://doi.org/10.1212/wnl.40.8.1218.
[9] S. Li, B. Ritz, Y. Gong, M. Cockburn, A.D. Folle, I. Del Rosario, Y. Yu, K. Zhang, E. Castro, A.M. Keener, J. Bronstein, K.C. Paul, Proximity to residential and workplace pesticides application and the risk of progression of Parkinson's diseases in Central California, Sci. Total Environ. 864 (2023), https://doi.org/10.1016/j.scitotenv.2022.160851.
[10] B. Ritz, S. Costello, Geographic model and biomarker-derived measures of pesticide exposure and Parkinson's disease, Ann. N. Y. Acad. Sci. 1076 (2006) 378–387, https://doi.org/10.1196/annals.1371.074.
[11] S. Costello, M. Cockburn, J. Bronstein, X. Zhang, B. Ritz, Parkinson's disease and residential exposure to maneb and paraquat from agricultural applications in the central valley of California, Am. J. Epidemiol. 169 (2009) 919–926, https://doi.org/10.1093/aje/kwp006.
[12] B. Krzyzanowski, Nielsen S. Searles, J.R. Turner, B.A. Racette, Fine particulate matter and Parkinson disease risk among Medicare beneficiaries, Neurology 101 (2023) e2058–e2067, https://doi.org/10.1212/WNL.0000000000207871.
[13] A.G. Dexter, Herbicide Spray Drift, North Dakota State University Extension Service, Fargo, ND, 1993, p. 11. https://library.ndsu.edu/ir/bitstream/handle/10365/3067/126dex93.pdf?sequence=1.
[14] I. Faust, M. Warden, A. Camacho-Soto, B.A. Racette, S. Searles Nielsen, A predictive algorithm to identify ever smoking in medical claims-based epidemiologic studies, Ann. Epidemiol. 85 (2023) 59–67.e56, https://doi.org/10.1016/j.annepidem.2023.04.019.
[15] D.A. Goolsby, E.M. Thurman, M.L. Pomes, M.T. Meyer, W.A. Battaglin, Herbicides and their metabolites in rainfall: origin, transport, and deposition patterns across the midwestern and northeastern United States, 1990-1991, Environ. Sci. Technol. 31 (1997) 1325–1333, https://doi.org/10.1021/es960847o.
[16] B.R. Ritz, K.C. Paul, J.M. Bronstein, Of pesticides and men: a California story of genes and environment in Parkinson's disease, Curr Environ Health Rep 3 (2016) 40–52, https://doi.org/10.1007/s40572-016-0083-2.
[17] Q. Yan, K.C. Paul, D.I. Walker, M.A. Furlong, I. Del Rosario, Y. Yu, K. Zhang, M. G. Cockburn, D.P. Jones, B.R. Ritz, High-resolution metabolomic assessment of pesticide exposure in Central Valley, California, Chem. Res. Toxicol. 34 (2021) 1337–1347, https://doi.org/10.1021/acs.chemrestox.0c00523.
[18] Merlos F. Aramburu, R.J. Hijmans, The scale dependency of spatial crop species diversity and its relation to temporal diversity, Proc. Natl. Acad. Sci. USA 117 (2020) 26176–26182, https://doi.org/10.1073/pnas.2011702117.
[19] J. Aguilar, G.G. Gramig, J.R. Hendrickson, D.W. Archer, F. Forcella, M.A. Liebig, Crop species diversity changes in the United States: 1978-2012, PLoS One 10 (2015) e0136580, https://doi.org/10.1371/journal.pone.0136580.
[20] J. Knodel, A. Varenhorst, P. Beauzay, M. Boetel, A. Thostenson, Chlorpyrifos insecticide products legal for insect control in 2024, in: K. Anderson (Ed.), North Dakota Department of Agriculture, Fargo, ND, 2024, p. 1. https://www.ag.ndsu.edu/news/newsreleases/2024/march/chlorpyrifos-insecticide-products-legal-for-insect-control-in-2024.
[21] A.F. Wiese, W.L. Harman, B.W. Bean, C.D. Salisbury, Effectiveness and economics of dryland conservation tillage systems in the Southern Great Plains, Agron. J. 86 (1994) 725–730, https://doi.org/10.2134/agronj1994.00021962008600040026x.
[22] P.B. Beauzay, J.J. Knodel, G.A.S.M. Ganehiarachchi, Integrated Pest Management of Alfalfa Weevil in North Dakota, North Dakota State University Extension, Fargo, ND, 2018, p. 8. https://www.ndsu.edu/agriculture/sites/default/files/2022-08/e1676.pdf.
[23] J.B. Johnson, W.E. Zidack, S.M. Capalbo, J.M. Antle, D.F. Webb, Pests, Pesticide Use, and Pesticide Costs on Larger Central and Eastern Montana Farms with Annually-Planted Dryland Crops, Montana State University, Bozeman, MT, 1997, p. 22. https://www.montana.edu/extensionecon/pdf/sr23-1.pdf.
[24] V. Kumar, R. Liu, A.J. Jhala, P. Jha, M. Manuchehri, Palmer amaranth (Amaranthus palmeri) control in postharvest wheat stubble in the Central Great Plains, Weed Technol. 35 (2021) 945–949, https://doi.org/10.1017/wet.2021.64.
[25] S. Smarik, S. Aney, A. Boes, D. Brown, D. Dubois, B. Edwards, E. Elias, M. Eve, R. Steele, N. Webb, M. Wilson, G. Zwicke, Understanding Wind Erosion; Soil Disturbance and Wind Erosion Processes. Dust Mitigation Handbook, U.S. Department of Agriculture, Las Cruces, NM, 2019, pp. 7–14. https://dust.swclimatehub.info/2.1.
[26] F.M. Fishel, J.A. Ferrell, Managing pesticide drift. University of Florida - Institute of Food and Agricultural Sciences, FL, Gainesville, 2019, p. 15. https://edis.ifas.ufl.edu/publication/PI232.
[27] G.A. Matthews, Pesticide application in relation to integrated pest management, Int. J. Trop. Insect Sci. 15 (1994) 599–604, https://doi.org/10.1017/S1742758400016908.
[28] N.C. Hansen, B.L. Allen, R.L. Baumhardt, D.J. Lyon, Research achievements and adoption of no-till, dryland cropping in the semi-arid US Great Plains, Field Crops Res. 132 (2012) 196–203, https://doi.org/10.1016/j.fcr.2012.02.021.
[29] Y. Hu, C. Zhou, C. Tan, J. Liu, X. Huang, X. Liu, C. Yao, D. Li, Q. Huang, N. Li, J. Long, X. Li, L. Yafei, L. Zhou, T. Cai, The association between intermediate-term sulfur dioxide exposure and outpatient visits for Parkinson's disease: a time-series study in southwestern China, Environ. Sci. Pollut. Res. Int. 30 (2023) 99694–99703, https://doi.org/10.1007/s11356-023-29408-6.
[30] R. Liu, M.T. Young, J.C. Chen, J.D. Kaufman, H. Chen, Ambient air pollution exposures and risk of Parkinson disease, Environ. Health Perspect. 124 (2016) 1759–1765, https://doi.org/10.1289/EHP135.