

## Linking pesticides and human health: A geographic information system (GIS) and Landsat remote sensing method to estimate agricultural pesticide exposure



Trang VoPham <sup>a,\*</sup>, John P. Wilson <sup>b</sup>, Darren Ruddell <sup>b</sup>, Tarek Rashed <sup>b</sup>, Maria M. Brooks <sup>a</sup>, Jian-Min Yuan <sup>a,c</sup>, Evelyn O. Talbott <sup>a</sup>, Chung-Chou H. Chang <sup>d</sup>, Joel L. Weissfeld <sup>a,c</sup>

<sup>a</sup> Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, 130 De Soto St, Pittsburgh, PA 15261, USA

<sup>b</sup> Spatial Sciences Institute, University of Southern California, 3616 Trousdale Pkwy AHF B55, Los Angeles, CA 90089, USA

<sup>c</sup> Division of Cancer Control and Population Sciences, University of Pittsburgh Cancer Institute, 5150 Centre Ave, Pittsburgh, PA 15232, USA

<sup>d</sup> Department of Medicine, University of Pittsburgh, 1218 Scaife Hall, 3550 Terrace St, Pittsburgh, PA 15261, USA

### ARTICLE INFO

#### Article history:

Available online 18 May 2015

#### Keywords:

Pesticide exposure  
Geographic information system (GIS)  
Remote sensing  
Normalized Difference Vegetation Index (NDVI)  
Environmental epidemiology

### ABSTRACT

Accurate pesticide exposure estimation is integral to epidemiologic studies elucidating the role of pesticides in human health. Humans can be exposed to pesticides via residential proximity to agricultural pesticide applications (drift). We present an improved geographic information system (GIS) and remote sensing method, the Landsat method, to estimate agricultural pesticide exposure through matching pesticide applications to crops classified from temporally concurrent Landsat satellite remote sensing images in California. The image classification method utilizes Normalized Difference Vegetation Index (NDVI) values in a combined maximum likelihood classification and per-field (using segments) approach. Pesticide exposure is estimated according to pesticide-treated crop fields intersecting 500 m buffers around geocoded locations (e.g., residences) in a GIS. Study results demonstrate that the Landsat method can improve GIS-based pesticide exposure estimation by matching more pesticide applications to crops (especially temporary crops) classified using temporally concurrent Landsat images compared to the standard method that relies on infrequently updated land use survey (LUS) crop data. The Landsat method can be used in epidemiologic studies to reconstruct past individual-level exposure to specific pesticides according to where individuals are located.

© 2015 Elsevier Ltd. All rights reserved.

### Introduction

Pesticides, chemicals designed to treat pests such as insects, have been associated with adverse human health outcomes such as cancers (Alavanja, Hoppin, & Kamel, 2004; Blair, Ritz, Wesseling, & Beane Freeman, 2015). One source through which pesticide exposure may impact human health is via residential

proximity to agricultural pesticide applications (Rull & Ritz, 2003; Ward et al., 2000). Applied pesticides may drift through the air and the ground and through post-application volatilization. Large-scale pesticide drift incidents frequently occur in agricultural areas in California (CA), United States (US), impacting residents and field workers and resulting in acute symptoms such as vomiting and impaired breathing (Harrison, 2006). In California, upwards of 90%

*Abbreviations:* AI, active ingredient; CA, California; CCM, compressed county mosaic; CDPR, California Department of Pesticide Regulation; CDWR, California Department of Water Resources; COST, cosine estimation of atmospheric transmittance; DNR, Department of Natural Resources; GIS, geographic information system; LUS, land use survey; MLC, maximum likelihood classification; NAD83, North American Datum 1983; NAIP, National Agriculture Imagery Program; NASA, National Aeronautics and Space Administration; NDVI, Normalized Difference Vegetation Index; NIR, near-infrared; NPS, nonpoint source; PLSS, Public Land Survey System; PUR, Pesticide Use Report; PUS, Pesticide Usage Survey; R, red (spectral band); SPOT, Satellite Pour l'Observation de la Terre; SRS, stratified random sampling; TM, Thematic Mapper; US, United States; USDA, United States Department of Agriculture; USGS, United States Geological Survey.

\* Corresponding author. Present address: Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 401 Park Dr, Boston, MA 02215, USA. Tel.: +1 757 478 2515.

E-mail addresses: [vophamt@gmail.com](mailto:vophamt@gmail.com) (T. VoPham), [jpwilson@dornsife.usc.edu](mailto:jpwilson@dornsife.usc.edu) (J.P. Wilson), [druddell@usc.edu](mailto:druddell@usc.edu) (D. Ruddell), [rashed@gsareh.com](mailto:rashed@gsareh.com) (T. Rashed), [brooks@edc.pitt.edu](mailto:brooks@edc.pitt.edu) (M.M. Brooks), [yuanj@upmc.edu](mailto:yuanj@upmc.edu) (J.-M. Yuan), [eot1@pitt.edu](mailto:eot1@pitt.edu) (E.O. Talbott), [changj@pitt.edu](mailto:changj@pitt.edu) (C.-C.H. Chang), [jwepid@comcast.net](mailto:jwepid@comcast.net) (J.L. Weissfeld).

of registered pesticide products are prone to drift. Gunier et al. (2011) demonstrated that pesticides measured in carpet dust from 89 residences in California were significantly correlated with residential proximity to agricultural pesticide applications quantified using a geographic information system (GIS) (Spearman correlation coefficients 0.23 to 0.50;  $p < 0.05$ ). Humans are subsequently affected by pesticides through dermal contact and ingestion, especially as pesticides are less likely to degrade within houses (Gunier, Harnly, Reynolds, Hertz, & Von Behren, 2001).

Elucidating the exact role pesticide exposure may play in the risk of developing adverse health outcomes is impacted by the methods used to quantify exposure. GIS metrics can combine multiple data sources to reconstruct historical exposure to specific pesticides (Franklin & Worgan, 2005). The California Department of Pesticide Regulation (CDPR) has collected Pesticide Use Report (PUR) data pertaining to agricultural use pesticide applications since 1974, including pounds (1 pound represents 0.45 kg) of pesticides used to treat specified crop types on specified dates within Public Land Survey System (PLSS) sections (CDPR, 2014). However, PUR data alone cannot be used to match pesticide applications to specific geographic locations at a scale finer than the 2.59 km<sup>2</sup> (1 mi<sup>2</sup>) PLSS section level. This limitation has motivated attempts to combine PURs with land use data, notably the California Department of Water Resources (CDWR) land use surveys (LUS's). Rull and Ritz (2003) developed the standard validated GIS method of estimating agricultural pesticide exposure in California via a three-tier methodology that assigns PUR pounds of applied pesticides to LUS crop fields (Rull, Ritz, & Shaw, 2006a, 2006b). The notion of “tiers” refers to the level of certainty with which a PUR pesticide application can be assigned to a particular LUS crop field. Combining PURs with a LUS enables pesticide application rate calculations at geographic scales finer than the PLSS level. However, CDWR LUS's are infrequently conducted on a county basis once every seven to 10 years, during which time significant land use changes can occur (Nuckols et al., 2007).

Although vector data have typically dominated this research, raster data provide a valuable way to incorporate temporally concurrent land use information in pesticide exposure estimation. Ward et al. (2000) pioneered the integration of Landsat remote sensing, which has continuously captured satellite imagery of the Earth since 1972 (USGS, 2014), in estimating pesticide exposure. Supervised classification of a Landsat image of Nebraska, US from 1984 was implemented to classify agricultural land cover types, which were subsequently assigned crop-specific pesticide use probabilities. Wan (2015) developed a GIS and remote sensing method to estimate population-level exposure using Nebraska land use data (classified from Landsat images), United States Geological Survey (USGS) county-level pesticide data, and crop-specific pesticide usage from farmer surveys. Population density grid cells were assigned pesticide exposure values according to downscaled pesticide data using 1,000 m radius buffers around cell centroids. Maxwell, Airola, and Nuckols (2010) demonstrated how Landsat imagery of California could be used to downscale the identification of PUR pesticide-treated crop fields below the LUS level (minimum mapping unit 0.008 km<sup>2</sup>) (Nuckols et al., 2007). Normalized Difference Vegetation Index (NDVI) values, a measure of vegetative density, were used to classify imagery into crop fields via a minimum distance method, and when used in conjunction with PLSS sections, can identify probable crops treated with pesticides (Maxwell, 2011).

However, minimum distance classification is not widely used in practice as it cannot take into account the spectral variability present within land use classes (Campbell & Wynne, 2011). Alternative approaches include implementing per-pixel maximum likelihood classification (MLC) using NDVI values (De Wit & Clevers, 2004;

Guerschman, Paruelo, Bella, Giallorenzi, & Pacin, 2003) and/or per-field classification, which is useful in addressing within-field spectral heterogeneity (Lu & Weng, 2007). For example, Turker and Ozdarici (2011) implemented per-field classification, where ML-classified pixels of SPOT, IKONOS, and QuickBird imagery of Turkey in 2004 were used to classify vector fields according to the most frequently occurring land use class pixel.

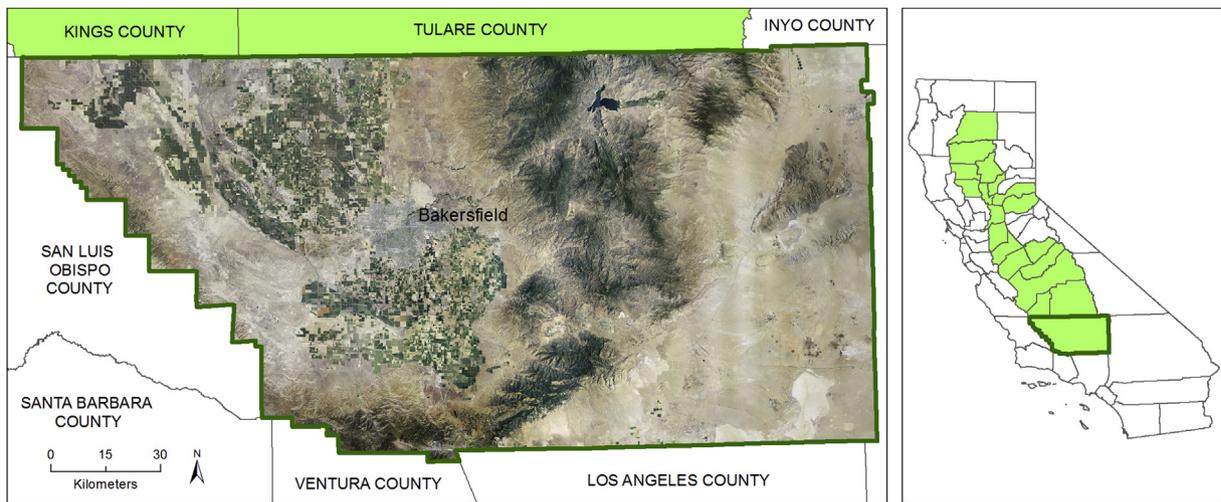
This study demonstrates the use of an improved GIS and remote sensing method, the Landsat method, to estimate agricultural pesticide exposure in a year without a concurrent standard LUS crop field dataset. The Landsat method matches PUR pesticide application data to concurrent Landsat images that have been classified into crops via an MLC and per-field classification approach. Pesticide-treated crop fields intersecting 500 m buffers around geocoded locations are used to estimate pesticide exposure. Our first research objective was to execute an accuracy assessment comparing classified Landsat images in 1990 to the 1990 LUS gold standard (ground truth). As part of this first objective, we determined the accuracy of 1990 agricultural pesticide exposure estimates using classified Landsat images from 1990 vs. the 1990 LUS. Our second research objective was to evaluate the crop specificity of 1985 pesticide applications matched to classified Landsat images, a demonstration of the Landsat method's utility. As part of this second objective, we compared pesticide exposure estimates derived from 1985 pesticide application data matched to classified Landsat images from 1985 vs. the 1990 LUS.

## Methods

### *Study area and data sources*

Kern County, CA, US is 21,061.58 km<sup>2</sup> in area and is one of 19 counties in the agriculturally intensive Central Valley (Fig. 1) (USDA, 2003). Agricultural croplands are predominantly found in the central and northwestern portions of the county. From 1982 to 1992, the majority of Kern County's farm area (4,058–3,900 km<sup>2</sup>) was associated with harvested cropland (76.6–86.7%), which was consistently dominated by cotton (34.6–36.9%) (USDA, 2014).

The CDPR PURs include California agricultural pesticide application data from 1974 to present (full use reporting started in 1990) (CDPR, 2014). PUR data include the name, pounds (1 pound represents 0.45 kg), date, crop, and PLSS section associated with reported pesticide applications. The PLSS divides portions of the US into 2.59 km<sup>2</sup> (1 mi<sup>2</sup>) sections, each identified by a county, principal meridian, township, range, and section (National Atlas, 2014). USGS and National Aeronautics and Space Administration (NASA) Landsat satellites have collected Earth imagery since 1972 (USGS, 2014). The Thematic Mapper (TM) sensor onboard Landsat 4 and 5 (used in this analysis) captured seven spectral bands with at least 30 m spatial resolution. Each Landsat scene, defined by a Path–Row designation, spans 185 km and is captured every 16 days. Bands 3 (red; 0.63–0.69 μm) and 4 (near-infrared; 0.76–0.90 μm) were used in this analysis to calculate NDVI values (Maxwell, Airola, et al., 2010; Maxwell, Meliker, & Goovaerts, 2010), which harness information from wavelengths of electromagnetic radiation absorbed and reflected by green plants and their changes throughout the growing season (USGS, 2011). NDVI values range from –1 (no or sparse vegetation) to 1 (dense vegetation). The CDWR conducts LUS's of agricultural lands to monitor land use changes in California on a county basis focusing on over 70 crop types (CDWR, 2014). Each LUS dataset is updated every seven to 10 years. Residential parcels were selected from the 2012 Kern County Assessor file via use codes (e.g., 0100, single family residence) (Kern County Assessor, 2012). All administrative boundaries were mapped



**Fig. 1.** Kern County, CA, US study area. A National Agriculture Imagery Program (NAIP) compressed county mosaic (CCM) of Kern County from August 2012 is shown on the left (USDA Farm Service Agency Aerial Photography Field Office, 2012). Kern County within California's Central Valley agricultural region is shown on the right.

using United States Census Bureau TIGER/Line© files (U.S. Census Bureau, 2014).

#### Implementing the Landsat method

The ultimate goal of the Landsat method is to estimate pesticide exposure experienced by individuals according to the pesticide-treated crop fields near their residences. The Landsat method requires four pieces of information: (1) Landsat satellite images, (2) ground truth, (3) pesticide application data, and (4) geocoded locations (e.g., residences). Landsat images of Kern County, a LUS ground truth, California PUR pesticide data, and residential parcels were used for this study. The Landsat method is comprised of three main steps: (1) classifying monthly Landsat NDVI images into crops, (2) matching pesticide applications to the classified Landsat crops, and (3) estimating pesticide exposure according to pesticide-treated Landsat crops intersecting 500 m (radius) buffers around geocoded locations in a GIS. Five hundred meters corresponds to a relevant distance within which to estimate human pesticide exposure previously used in epidemiologic studies (Rull & Ritz, 2003; Rull et al., 2006a). Aerial pesticide applications can drift between 500 and 1,000 m from where they were applied (Ward et al., 2000).

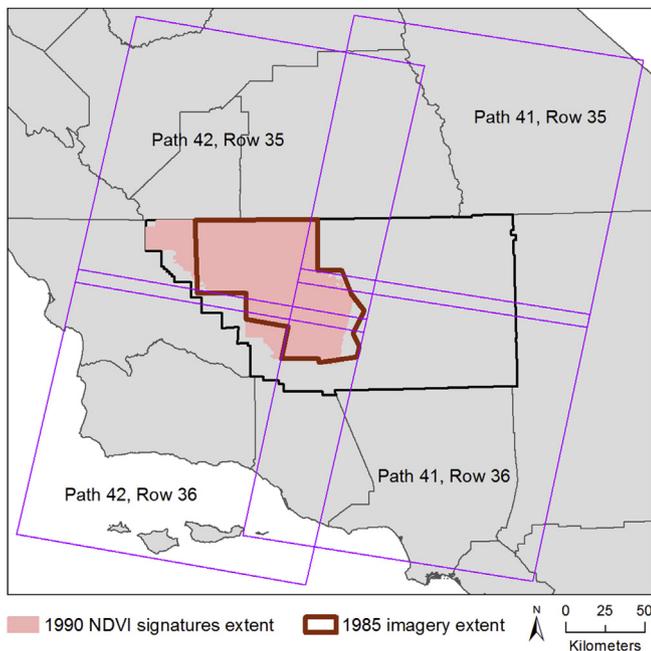
NDVI images, which show vegetative density across the study area, are created using Landsat images and classified into crops using a supervised, hard, per-pixel MLC and per-field (using segments) classification approach that assigns one particular crop type to each segment for the time period of analysis. A monthly time series of cloud-free images is used to enhance land use discrimination. In order to classify the NDVI images, training data must be created, which are comprised of monthly NDVI images and a ground truth. When implementing this method in practice, training data should be created (1) using data in the geographic study area of interest and (2) in a year close in time to the year pesticide exposure is to be estimated to minimize spatiotemporal changes/differences in land uses, growing practices, Landsat sensors, and NDVI values. Furthermore, it is very important to perform an accuracy assessment in the year used to create training data to examine land use classification accuracies. In California, PUR pesticide applications from any year between 1974 to present day can be matched to Landsat images classified into crops from that same year.

#### Landsat image and LUS preprocessing

The year of 1990 was selected for the accuracy assessment as there was an available Kern County LUS ground truth and Landsat images. The year of 1985 was selected to demonstrate pesticide exposure estimation as a concurrent LUS is not available and it was the first year moving backward from 1990 that had available cloud-free images from January to October paralleling the monthly images used to create the 1990 training data. A time series of Landsat 4 and 5 TM monthly images captured between January and October 1990 (no available November and December images) and January and October 1985 was downloaded (Supplemental Tables S1 and S2). Images from Paths 41 and 42 and Rows 35 and 36 were requested, which cover the geographic extent of Kern County (Fig. 2) (USGS, 2013). Portions of images with excessive cloud cover were excluded. Portions of the February 1985 image missing Path 42 (majority of Kern County agricultural fields) were imputed with the average of the January and March 1985 images (Martinuzzi, Gould, & Ramos González, 2007). Using IDRISI Selva (Clark Labs, 2014), images for the red (R) and near-infrared (NIR) bands (used to calculate NDVI) were corrected to at-sensor reflectance (Chander, Markham, & Helder, 2009). Atmospheric correction was implemented via the Chavez cosine estimation of atmospheric transmittance (COST) model (Chavez, 1996; Song, Woodcock, Seto, Lenney, & Macomber, 2001). Path–Row images were mosaicked and negative reflectance values were recoded to 0 (Yale Center for Earth Observation, 2013). A median spatial filter (3x3 kernel) was applied to each mosaicked image (Vassiliou, Boulianne, & Blais, 1988). NDVI values were calculated using the following equation:  $\text{NDVI} = \frac{\text{NIR} - \text{R}}{\text{NIR} + \text{R}}$ . NDVI images from 1990 were cropped to what is referred to as the 1990 NDVI signatures extent in Fig. 2, a region defined by images unaffected by clouds and/or shadows and within the 1990 Kern County LUS surveyed area. NDVI images from 1985 were cropped to what is referred to as the 1985 imagery extent in Fig. 2, a cloud- and shadow-free area within the 1990 Kern County LUS extent. NDVI images were re-projected to the California Teale Albers (NAD83 datum; meter) coordinate system (30 m spatial resolution; nearest neighbor resampling to not alter pixels).

#### Creation of training data and classification of Landsat images

Polygons of single-use (e.g., excluding double-cropped), representing classified areas (e.g., excluding outside study area [Z]), and



**Fig. 2.** Landsat Path-Rows intersecting Kern County, CA and geographic extent of NDVI training data and classification data. Landsat scenes were requested for these four Path–Row combinations. NDVI signatures for land use classes derived from images within the 1990 NDVI signatures extent (pink region) were used for the accuracy assessment in 1990 and as training data for MLC of 1985 images within the 1985 imagery extent (red region). For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

within the 1990 NDVI signatures extent were selected from the 1990 Kern County LUS. A negative buffer (–30 m; spatial resolution of Landsat images) was created around each selected LUS polygon to exclude potential mixed pixels (CDWR, 2009). After collapsing urban LUS polygons into a single category, LUS polygons with valid geometries were intersected with the NDVI images from 1990. Land uses represented by fewer than 100 pixels in each month in 1990 were excluded from consideration for training data in order to include representative classes (Richards, 2013), resulting in 57 distinct land use classes in the training data.

Stratified random sampling (SRS) selected 60% of the buffered polygons from the 1990 LUS within the 1990 NDVI signatures extent to be used as training data in the accuracy assessment (CDWR, 2009). Strata were defined by the land use classes. The remaining 40% of the 1990 NDVI signatures extent, referred to as the 40% classification extent, was segmented and classified using MLC and per-field classification. Using ArcGIS 10.1 (Esri, 2014), MLC and the sample option assigned a priori probabilities to land use classes in proportion to the number of cells represented in the training data (Mueller-Warrant et al., 2011). Using IDRISI Selva, segmentation was performed on the monthly 1990 NDVI images within the 40% classification extent using the following parameters: window of 3, tolerance of 0.01, weight mean factor of 0.5, and weight variance factor of 0.5. Using the ML-classified pixels, per-field classification (using the segments) was implemented based on the modal class or a majority rule (Turker & Ozdarici, 2011; Van Niel & McVicar, 2004). For the 1985 analysis, all training data from the 1990 NDVI signatures extent, not restricted the 60% SRS, were used to classify the monthly 1985 NDVI images using MLC and per-field classification.

#### Pesticide exposure estimation

The three-tier method (Rull & Ritz, 2003) was implemented to estimate pesticide application rates in 1990 and 1985 by matching

PUR pesticide applications to either classified 1990 or 1985 Landsat images (referred to as the Landsat method) or the 1990 Kern County LUS (referred to as the LUS method; LUS conducted closest in time to the 1985 PUR data). PUR data were processed using CDPR logic checks such as duplicate removal (CDPR, 2014). Outlier application rates in 1990 were defined using CDPR-created flags (CDPR, 2014) and in 1985 as pesticide application rates  $>22,417$  kg/km<sup>2</sup> ( $>112,085$  kg/km<sup>2</sup> if fumigant) or pesticide application rates greater than 50 times the median rate for all uses of a given pesticide product, crop, unit type, and record type. Outliers were replaced with the statewide median rate for the pesticide active ingredient (AI) in that year. Pounds of AI were recalculated using the number of treated acres (Rull & Ritz, 2003). Organophosphates were identified using agricultural references (AgroPages, 2014; Alavanja et al., 2004; Dich, Zahm, Hanberg, & Adami, 1997; Greene & Pohanish, 2005; Gunier et al., 2001; Rull et al., 2009, 2006a; Wood, 2010). A crosswalk between PUR crop codes and CDWR LUS crop codes was created to facilitate data linkage.

Landsat method crop fields used to match to PUR pesticide applications were derived from segments classified as agricultural use (e.g., grain) that were spatially joined to the 2,337 PLSS sections intersecting the 40% classification extent (accuracy assessment) or the 2,491 PLSS sections within the 1985 imagery extent. Segments were dissolved according to crop type and section, the geographic level of reporting of the PUR database. LUS method crop fields used to match to PURs were derived from agricultural use LUS polygons selected from the 40% classification extent or within the 1985 imagery extent, spatially joined to sections, and dissolved according to crop type and section. The three-tier method hierarchically matched pesticide applications to one of three tiers. Tier 1: Pesticides were matched to a crop field according to crop type and section. Tier 2: Pesticides were matched to all other crop fields in a section. Tier 3: Pesticides were matched to the entire section. The LUS method implemented in this study did not collapse non-permanent crop fields into a single category to facilitate the examination of specific crop types. Using the tier-matched organophosphates, pesticide application rates (kg/km<sup>2</sup>) were calculated for sampled residential parcels separately using the Landsat and LUS methods. SRS selected at most three residential parcel centroids from each of the sections (strata) within the 40% classification extent or the 1985 imagery extent, and 500 m (radius) buffers were created around the centroids of sampled residential parcels. Area estimates were derived from Landsat, LUS, and section data. Pesticide application rates were weighted by the proportion of pesticide-treated crop fields and/or sections intersecting the buffer.

#### Statistical analysis and error matrices

For the accuracy assessment, classified Landsat segments and the LUS were intersected and compared by segment and total area using site-specific error matrices. Agreement, kappa and 95% confidence intervals (CIs), producer's accuracy, user's accuracy, omission error, and commission error were calculated according to CDWR land use (i.e., specific crop), CDWR broad land use group (e.g., field crops), and phenological group. Phenological groups were determined using the SAS Proc Cluster centroid method, which grouped together land use classes based on the squared Euclidean distance between their centroids (mean NDVI value for each land use class for each month) (Guerschman et al., 2003; Simoniello, Carone, Lanfredi, Macchiato, & Cuomo, 2004; Suzuki, Nomaki, & Yasunari, 2001). Bowker's test of symmetry for paired data compared the proportion of tier 1, tier 2, and tier 3 matches when using the Landsat vs. the LUS method. McNemar's tests compared the proportion of tier 1 vs. tier 2 and 3 matches, and tier 1 and 2 vs. tier 3 matches, when using the Landsat vs. LUS method,

as well as the proportion of tier 1 vs. tier 2 and 3 matches by crop type according to each method. Wilcoxon signed-rank tests compared pesticide application rates estimated using each method, and Spearman rank coefficients quantified the correlation between rates. Weighted kappa coefficients quantified the agreement in pesticide exposure categorizations according to each method. Analyses were conducted in 2014 using SAS version 9.4 (SAS Institute, Inc., Cary, North Carolina).

## Results

### Accuracy assessment in 1990: land use classification and pesticide exposure estimation

A total of 3,634.39 km<sup>2</sup> were used to create training data in 1990, representing the 60% stratified random sample of the 1990 NDVI signatures extent that in turn classified 2,532.36 km<sup>2</sup>. Landsat segments were on average 0.03 km<sup>2</sup> (median 0.02) in size, compared to LUS polygons that were on average 0.34 km<sup>2</sup> (median 0.20). There was an average of 29 (22 standard deviation [SD]) pixels (median 24) available to classify each segment.

When selecting the intersections between the segments and the single-use LUS polygons comprising the majority of each segment's original area, agreement was substantially high at the CDWR land use level (top row of Table 1) (Landis & Koch, 1977). The highest producer's accuracy was observed for asparagus (15/15=100%). The highest user's accuracy was observed for cotton (19,059/20,605=92.5%). Kappa statistics improved when aggregating segments and LUS polygons into CDWR broad land use groups and into phenological groups (Table 1). Select phenological groups out of a total of 17 representing NDVI patterns over a calendar-year time period are shown in Fig. 3. Comparable results were observed when examining the entire area of the intersections (bottom row of Table 1).

A closer examination of the accuracy assessment aggregated to CDWR broad land use groups reveals satisfactory producer's and user's accuracy for the majority of the agricultural broad land use groups (Table 2). Producer's accuracy was upwards of 82.6% for pasture crops and user's accuracy was upwards of 88.6% for field crops. Among agricultural broad land use groups, the highest omission error (96%) and commission error (91.9%) was observed for idle (i.e., fallow) lands. Truly idle lands were often misclassified as native vegetation (76.7%), while idle-classified segments were truly field crops (35.6%) or native vegetation (19.8%) - all of which belong to the same phenological group (Fig. 3). It is important to note that among segments that were classified as agricultural use, a high proportion (73.9–98.9%) truly belongs to an agricultural broad land use group as opposed to a non-agricultural group (NV, NW, S, or U) according to the LUS (Table 2).

According to 7,495 PUR applications in 1990, LUS (median 44.83 kg/km<sup>2</sup>; interquartile range [IQR] 0–195.03) and Landsat (median 51.56 kg/km<sup>2</sup>; IQR 0–210.72) pesticide application rates in 1990 were not significantly different for the 1,291 sampled residential parcels (Wilcoxon signed-rank  $p=0.8513$ ). Rates were significantly correlated (Spearman correlation 0.83;  $p<0.0001$ ). A similar number of crop types were present within any given section when using the LUS (mean 1.3; median 1.0) and Landsat layers (mean 2.8; median 2.0). A similar number of pesticide-treated crop types intersected any given 500 m buffer when using the LUS (mean 1.4; median 1.0) and Landsat layers (mean 1.8; median 1.0). Using quartiles defined by the distribution of LUS pesticide application rates (none: 0 kg/km<sup>2</sup>; low: >0–44.83 kg/km<sup>2</sup>; moderate: 44.83–195.03 kg/km<sup>2</sup>; high: >195.03 kg/km<sup>2</sup>), agreement between LUS and Landsat pesticide exposure classifications was high (weighted kappa 0.766, 95% CI 0.739, 0.792).

### Demonstration of Landsat method in 1985

A total of 50,441 segments (mean 0.14 km<sup>2</sup>; median 0.11 km<sup>2</sup>) derived from 1985 Landsat imagery were classified (Fig. 4). Fig. 5 shows the spatial relationships between PLSS sections, LUS crops, and segments prior to dissolving by crop type and section. There was an average of 153 (127 SD) pixels (median 120) available to classify each segment. The majority of segments were classified as alfalfa (19.5%), followed by cotton (19.3%), field crop (18.7%), and native vegetation (8.6%).

According to 3,909 PUR applications in 1985, the proportion of tier matches were significantly different when using the Landsat method vs. the LUS method (Bowker's  $p<0.0001$ ; Table 3). The Landsat method achieved a significantly higher proportion of tier 1 matches (60.3%) compared to the LUS method (57.4%) (McNemar's  $p=0.0002$ ). The Landsat method (99.2%) achieved significantly more combined tier 1 and 2 matches vs. tier 3 matches compared to the LUS method (96.6%; McNemar's  $p<0.0001$ ). Among the 2,466 PUR applications associated with temporary crops (e.g., cotton), the Landsat method (65.4%) achieved a significantly higher proportion of tier 1 matches compared to the LUS method (52.4%; McNemar's  $p<0.0001$ ). Among the 1,443 PUR applications associated with permanent crops (e.g., oranges), the LUS method (66.0%) achieved a significantly higher proportion of tier 1 matches compared to the Landsat method (51.6%; McNemar's  $p<0.0001$ ).

A larger proportion of PUR applications associated with the following temporary crops were matched at tier 1 to Landsat compared to the LUS: alfalfa ( $n=468$ ; Landsat 98% vs. LUS 76%; McNemar's  $p<0.0001$ ), dry beans ( $n=75$ ; 67% vs. 7%;  $p<0.0001$ ), cotton ( $n=792$ ; 97% vs. 83%;  $p<0.0001$ ), and potatoes ( $n=300$ ; 65% vs. 51%;  $p=0.0001$ ). Assuming PUR data are accurate, a larger proportion of tier 1 matches among temporary crops is indicative of

**Table 1**  
Accuracy assessment of classified Landsat imagery vs. LUS in 1990.<sup>a</sup>

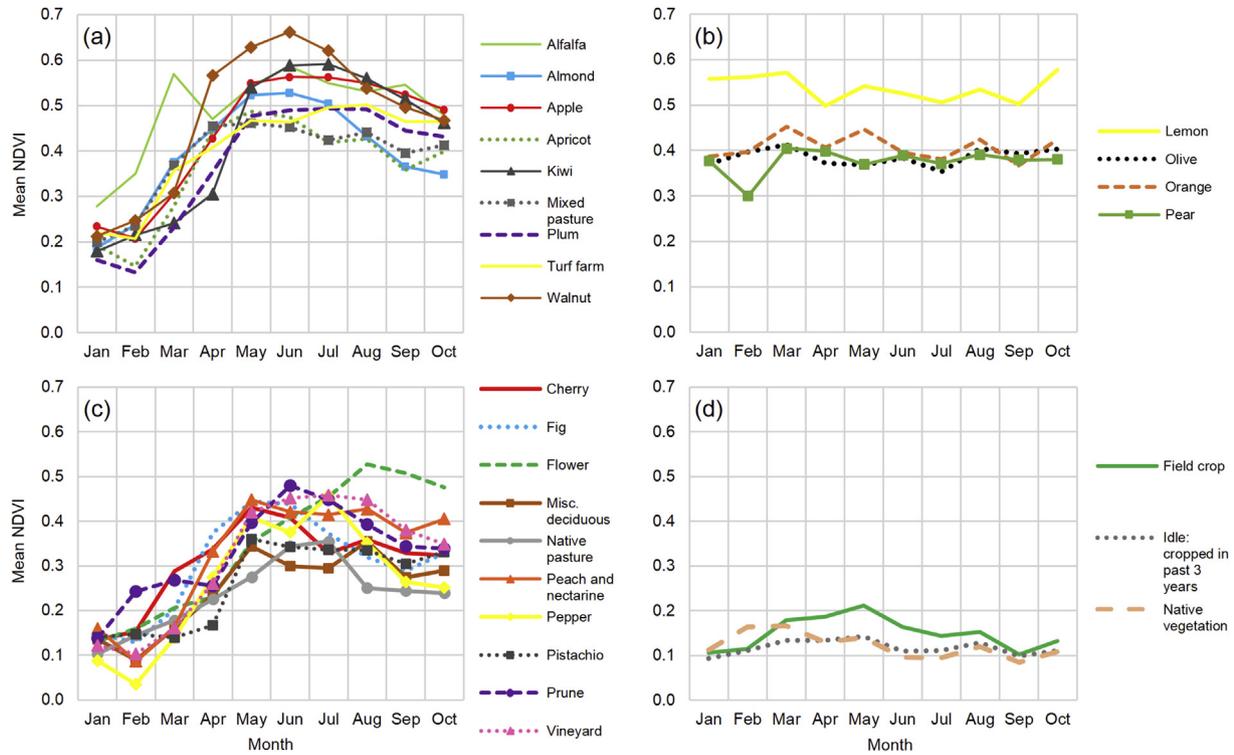
N	CDWR land use		CDWR broad land use group		Phenological group	
	Agreement	Kappa (95% CI)	Agreement	Kappa (95% CI)	Agreement	Kappa (95% CI)
87,197 segments <sup>b</sup>	75.1	0.700 (0.696, 0.703)	78.5	0.732 (0.728, 0.735)	82.7	0.779 (0.776, 0.782)
2,266 km <sup>2c</sup>	76.4	0.701 (0.699, 0.702)	79.5	0.731 (0.730, 0.733)	84.4	0.789 (0.788, 0.790)

Abbreviations: CDWR, California Department of Water Resources; CI, confidence interval; LUS, land use survey.

<sup>a</sup> The accuracy assessment was performed on classified imagery from the 40% classification extent that was derived from the 1990 NDVI signatures extent and not used for training data. Results are presented for the intersections between segments and single-use LUS polygons.

<sup>b</sup>  $n=86,060$  segments for phenological groups.

<sup>c</sup>  $n=2,243$  km<sup>2</sup> for phenological groups.



**Fig. 3.** Phenological groups comprised of land uses sharing similar annual NDVI patterns. These phenological groups, derived from a cluster analysis, include land uses that exhibit (a) a gradual summer NDVI peak, (b) a stable NDVI pattern, (c) a moderate vegetative density peak, and (d) a low NDVI pattern.

the capacity of Landsat imagery to delineate agricultural use lands not otherwise present in an outdated LUS. For example, Fig. 6 shows a residential parcel that was sampled in the 1985 imagery extent and located in section 15M29S25E10. PUR data indicated one organophosphate application of 12.68 kg occurred in PLSS section 15M29S25E15 (section south of parcel but intersecting parcel buffer) on alfalfa on October 5, 1985. No alfalfa fields were present in this section using the 1990 LUS (left), resulting in a tier 2 match and an estimated rate of 91.91 kg/km<sup>2</sup> for the selected residential parcel. However, 1985 Landsat imagery (right) identified alfalfa-

classified segments in this section, achieving a tier 1 match and an estimated rate of 128.90 kg/km<sup>2</sup>. The crop types present in section 15M29S25E15 according to the LUS and Landsat methods - alfalfa, cotton, and sugar beet - belong to three different phenological groups, providing support that the alfalfa-classified segments are not a result of phenological misclassification (i.e., misclassification of a segment as another land use belonging to the same phenological group).

PUR applications associated with the following permanent crops achieved more LUS vs. Landsat tier 1 matches: almonds

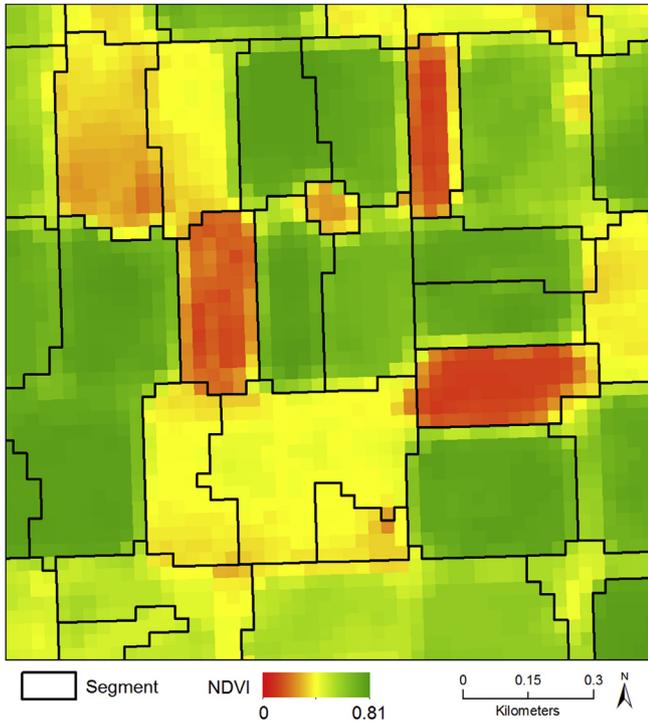
**Table 2**  
Accuracy assessment of Landsat imagery vs. LUS in 1990 using number of segments.<sup>a</sup>

	LUS (gold standard)													Total	User's	Agr <sup>b</sup>
	C	D	F	G	I	NV	NW	P	R	S	T	U	V			
C	<u>2,058</u>	44	17	9	5	76	1	8	0	19	23	76	42	2,378	86.5	92.8
D	27	<u>4,977</u>	42	7	33	291	22	190	0	30	23	203	456	6,301	79.0	91.3
F	68	201	<u>20,884</u>	312	111	517	24	499	0	35	628	50	242	23,571	88.6	97.3
G	10	17	466	<u>2,292</u>	65	364	3	32	0	2	228	4	0	3,483	65.8	89.3
I	13	63	406	91	<u>92</u>	226	6	58	0	33	97	33	24	1,142	8.1	73.9
NV	242	418	1,887	148	1,774	<u>22,770</u>	20	93	0	121	78	1,262	43	28,856	78.9	16.2
Landsat NW	0	2	16	1	2	53	<u>38</u>	0	0	0	2	10	0	124	30.6	18.5
P	48	136	775	47	89	29	0	<u>6,133</u>	0	10	81	45	176	7,569	81.0	98.9
R	0	0	0	0	0	0	0	0	<u>0</u>	0	0	0	0	0	—	—
S	1	2	1	0	1	9	0	3	0	<u>2</u>	0	1	3	23	8.7	47.8
T	6	20	985	139	52	133	0	195	0	1	<u>2,130</u>	7	31	3,699	57.6	96.2
U	95	203	262	28	63	571	10	56	0	174	<u>30</u>	<u>3,208</u>	94	4,794	66.9	17.3
V	33	583	164	7	27	230	6	161	0	25	91	<u>40</u>	<u>3,890</u>	5,257	74.0	94.3
Total	2,601	6,666	25,905	3,081	2,314	25,269	130	7,428	0	452	3,411	4,939	5,001	87,197		
Producer's	79.1	74.7	80.6	74.4	4.0	90.1	29.2	82.6	—	0.4	62.4	65.0	77.8			
Agr <sup>b</sup>	87.0	90.6	91.6	94.3	20.5	7.4	47.7	98.0	—	34.3	96.8	9.3	97.2			

Abbreviations: Agr, agricultural; C, citrus and subtropical; D, deciduous fruits and nuts; F, field; G, grain and hay; I, idle; LUS, land use survey; NV, native vegetation; NW, water surface; P, pasture; R, rice; S, semi-agricultural; T, truck, nursery, and berry; U, urban; V, vineyard.

<sup>a</sup> Segments and LUS polygons were aggregated into CDWR broad land use groups. Concordant cells have been underlined.

<sup>b</sup> Agricultural refers to the proportion of land use class classified as agricultural use (all classes except NV, NW, S, and U).



**Fig. 4.** Segments overlaying an August 1985 NDVI image. Segments were classified into land uses according to the most frequently occurring ML-classified pixel.

( $n=588$ ; LUS 85% vs. Landsat 75%; McNemar's  $p<0.0001$ ), oranges ( $n=322$ ; 91% vs. 75%;  $p<0.0001$ ), and peaches/nectarines ( $n=89$ ; 85% vs. 46%;  $p<0.0001$ ). Further examination of the crops associated with a greater number of LUS tier 1 matches revealed potential phenological misclassification. Among the 58 PLSS sections associated with LUS tier 1 almond matches, but no Landsat tier 1 matches, 95% contained segments classified as alfalfa, 26% with mixed pasture, and 12% with apples - all three of which belong to the same phenological group as almonds (Fig. 3).

LUS rates (median 20.18 kg/km<sup>2</sup>; IQR 0–67.25) were significantly different from Landsat rates (median 15.69 kg/km<sup>2</sup>; IQR 0–57.16) for the 1,293 sampled residential parcels (Wilcoxon

**Table 3**  
Pesticide application tier matching in 1985: LUS vs. Landsat methods.

		LUS <sup>a</sup>			Total	p <sup>b</sup>
		Tier 1	Tier 2	Tier 3		
Landsat <sup>a</sup>	Tier 1	1,864	452	40	2,356	<0.0001
	Tier 2	381	1,081	58	1,520	
	Tier 3	0	0	33	33	
	Total	2,245	1,533	131	3,909	

Abbreviations: LUS, land use survey.

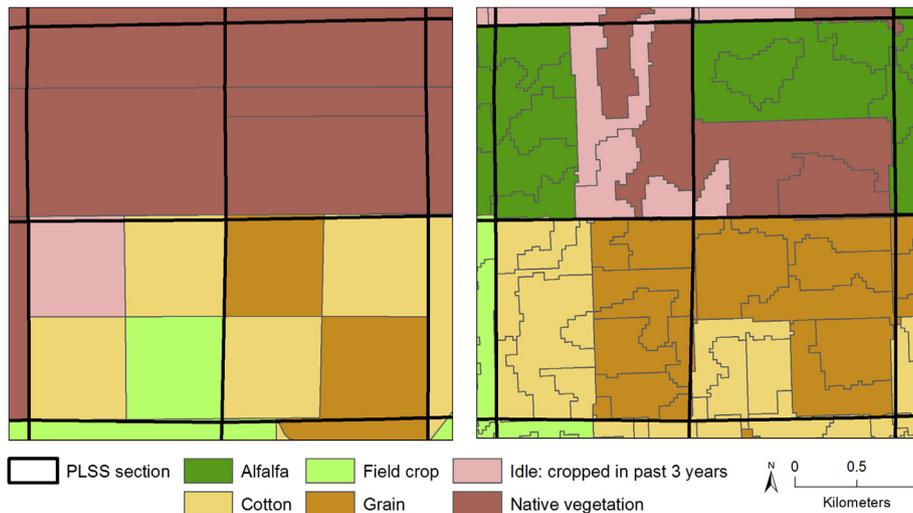
<sup>a</sup> 3,909 organophosphate applications (249,094.22 kg) in 1985 imagery extent.

<sup>b</sup> p-value reported from Bowker's test of symmetry.

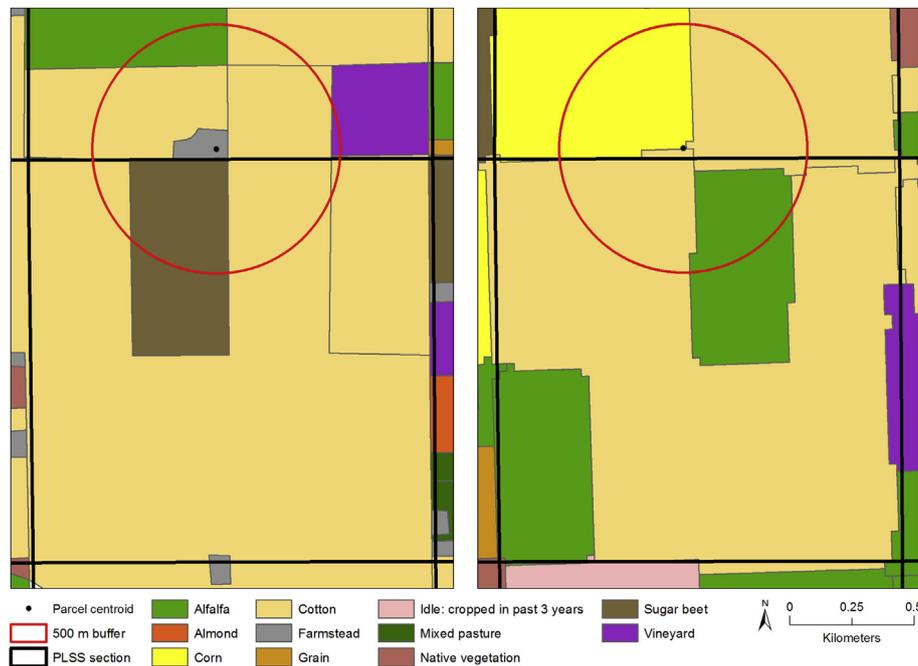
signed-rank  $p=0.0448$ ). However, a similar number of crops intersected any given section using the LUS (mean 3.1; median 3.0) and Landsat layers (mean 4.0; median 4.0). A similar number of pesticide-treated crops also intersected any given buffer using the LUS (mean 2.2; median 2.0) and Landsat layers (mean 2.7; median 2.0). Pesticide exposure classification according to LUS quartiles (none: 0 kg/km<sup>2</sup>; low: >0–20.18 kg/km<sup>2</sup>; moderate: 20.18–67.25 kg/km<sup>2</sup>; high: >67.25 kg/km<sup>2</sup>) demonstrated substantial agreement between both methods (weighted kappa 0.711, 95% CI 0.682, 0.740).

**Discussion**

GIS-based metrics are powerful tools in examining the relationship between pesticide exposure and human health outcomes. For example, the standard GIS method in California estimates agricultural pesticide exposure at residential locations through matching PUR pesticide applications with LUS crops (Rull & Ritz, 2003). However, dynamic agricultural landscapes, as a result of crop rotation and land use conversion (Chen, Li, & Allen, 2010), contribute to relevant changes that may impact GIS-based methods of estimating pesticide exposure. LUS's are intermittently updated every seven to 10 years on a county basis. Pesticide exposure estimation during a year lacking a temporally concurrent LUS will be affected as the utilized LUS may not adequately capture agricultural lands during that particular time period. Methods of incorporating remote sensing such as Landsat, which provide multispectral and multitemporal imagery capable of distinguishing landscape features (Maxwell, Meliker, & Goovaerts, 2010; USGS,



**Fig. 5.** PLSS sections overlaying 1990 Kern County LUS crops (left) and Landsat segments in 1985 (right). LUS crops and segments (created from monthly NDVI images in 1985) were each spatially joined to PLSS sections and dissolved according to crop type and section to facilitate PUR matching. For reference, this map includes section 15M27S20E01.



**Fig. 6.** Comparison of LUS vs. Landsat pesticide exposure estimation in 1985 for one residential parcel. One organophosphate alfalfa application (12.68 kg) in section 15M29S25E15 resulted in a tier 2 LUS match and LUS pesticide application rate of 91.91 kg/km<sup>2</sup> (left) and a tier 1 Landsat match and Landsat pesticide application rate of 128.90 kg/km<sup>2</sup> (right).

2014), allow for a useful approach to improving pesticide exposure estimation. The primary strengths of this research include the implementation of the Landsat method, an improved MLC and per-field classification approach to classify Landsat imagery into crops (compared to minimum distance methods), and the demonstration of a linkage between PUR data and Landsat-classified crops to estimate agricultural pesticide exposure in California in a year without a concurrent LUS.

The results of the accuracy assessment in 1990 provide evidence of the Landsat method's ability to both classify crops and estimate pesticide exposure as all results were in high agreement with the LUS gold standard method. The high correlation between Landsat and LUS pesticide application rates in 1990 bolsters the potentially negligible impact of any land use misclassification on pesticide exposure estimates. It should be noted that parcels from 2012 were used to estimate exposure in both 1990 and 1985. Although it is possible that sampled parcels were not present or were located in different areas during these times, any bias in exposure estimates would be nondifferential between the two methods. It is possible that the pesticide exposure accuracy results in 1990 may not reflect the accuracy of pesticide exposure estimation in 1985. A lower number of pesticide applications in 1985 vs. 1990 is a result of PUR reporting changes, increases in pesticide usage in California in the 1990s, pest outbreaks, and inclusion of different PLSS sections associated with the 40% classification extent vs. the 1985 imagery extent (CDPR, 2014; Godfrey, Rosenheim, & Goodell, 2000; Pesticide Action Network, 2013). However, as the proportion of the most commonly pesticide-treated crops were similar between 1985 and 1990 (20–30% cotton, 15% almonds, 12% alfalfa, 8% oranges, 7–8% table grapes), this lends confidence to the pesticide exposure accuracy assessment results generalizing to 1985 and little evidence of a differential proportion of treated crops between these two years.

The most prominent finding of the Landsat method demonstration in 1985 was the improvement in tier 1 matches, where significantly more pesticide applications were directly matched to Landsat crops vs. LUS crops - especially among temporary crops.

Temporary crops are sown/seeded and harvested during the same crop growing season (e.g., cotton), while permanent crops (trees [e.g., apples]) are sown or planted once and do not require replanting following harvests as they occupy the land for a long period of time (Food and Agriculture Organization, 2013). Assuming Landsat images in 1985 were accurately classified, the linkage between PUR data and Landsat crop fields was most beneficial to PUR applications associated with several truck, field, and pasture crops, which are comprised of temporary crops characterized by year-to-year changes. Agricultural growing practices are not limited to monocultures, or the repetitive growing of the same crop on the same land, but may include multiple cropping systems, also known as mixed cropping or polyculture, that intensify agricultural production through maximizing the efficiency of space and time to bio-diversify and stabilize the land, fertilize the crops in sequence, and promote pest control (Gliessman, 1985; Sullivan, 2003). For example, crop rotation consists of the repetitive growing of different crops in a systematic and recurring sequence on the same field, can be characterized by annual crop changes, and is widely practiced across the U.S (Liebman & Dyck, 1993; USDA, 2013).

The important theme underlying multiple cropping systems is the dynamic nature of agriculture, where any given year or growing season does not remain static. Using Landsat images to classify crops to match to PUR pesticide applications directly addresses these dynamic crop changes through not relying on outdated LUS's for crop fields. Promising results that serve to support the utility of a Landsat and PUR database linkage were demonstrated among the crops associated with Landsat tier 1 matches. A significantly higher proportion of 1985 organophosphate PUR applications associated with alfalfa, beans, cotton, and potatoes were able to be directly matched to 1985 Landsat-classified crops as opposed to 1990-dated LUS crops. These particular crops are temporary and are associated with documented crop rotation cycles in California, e.g., alfalfa is rotated with plants that do not host alfalfa-damaging pests (nematodes) such as cotton and beans (University of California Statewide Integrated Pest Management Program, 2006).

GIS and remote sensing methods of estimating human pesticide exposure apply concepts from nonpoint source (NPS) pollution spatial models of environmental risk assessment (Phillips, 1988). NPS pollution is the diffuse and dispersed contamination of surface or groundwater via runoff or leaching due to irrigation or precipitation. Our study of NPS pesticide pollution focused on pesticide drift and implemented NPS risk assessment concepts considering the spatial distribution and movement of pollutants, pollution sources, and affected resources. For example, this study used a 500 m buffer to address pesticide drift affecting residences, matched pesticide applications to crop fields to identify potential pollution sources, and used geocoded locations to identify potentially exposed individuals.

Although this study utilized California data sources, the Landsat method's framework, which requires Landsat images, ground truth, pesticide applications, and geocoded residences, can be applied to other U.S. states and countries. For example, the Iowa Department of Natural Resources (DNR) Land Cover datasets, capturing classes including corn and soybeans, could be used as ground truth to classify Landsat images (Natural Resources Geographic Information System Library, 2015). Iowa DNR interpolated density surfaces of yearly pesticide sales (lb sold per mi<sup>2</sup>) could be used as a proxy for applied pesticides. In the United Kingdom, the Pesticide Usage Survey (PUS) collects pesticide use from different sectors including agriculture and horticulture, and the Centre for Ecology and Hydrology creates Land Cover Maps that classify crops such as barley and carrots (Centre for Ecology and Hydrology, 2015; Fera, 2015). Matching of available pesticide data to agricultural lands may have to adopt methods beyond three-tier matching to accommodate discrepancies in spatial scales and in reported pesticide-treated crops such as aggregation or disaggregation. If studying a small geographic area, it may be feasible to acquire ground truth, pesticide applications, and geocoded residences via surveying methods.

#### *The Landsat method: implications for health and epidemiology*

GIS-based pesticide exposure metrics, such as the Landsat method, through their ability to incorporate multiple data sources with locational, dated information and specific chemicals, can address many important issues underlying the examination of chronic human health diseases including long latency periods (time between initial exposure and clinical diagnosis of disease), historical reconstruction to capture potential latency periods, multiple routes of exposure (e.g., dermal, inhalational, and oral), human exposure to multiple pesticides at different points in time, and recall bias (Franklin & Worgan, 2005). For example, the Landsat method can be implemented in any year in California with PUR pesticide data (1974 to present) and Landsat images (1972 to present) to estimate exposure to specific chemicals according to geocoded locations, capable of reconstructing past exposure relevant to epidemiologic studies seeking to address a latency period.

In an analytical epidemiologic study investigating if an exposure is associated with the risk of developing a particular disease, it is vital to provide the most accurate estimate of the exposure of interest. Otherwise, measures of association derived from the study are subject to bias. The Landsat method offers the opportunity to bridge the temporal gap between when pesticide exposure is to be estimated and the crop fields to which pesticide applications are matched. The Landsat method thus provides a way to minimize exposure misclassification when conducting an epidemiologic study based on findings from 1985 regarding more pesticide applications directly matching to Landsat crops compared to LUS crops. These tier 1 matches are ideal because they demonstrate a direct linkage between what is being reported in a PUR pesticide application in terms of the treated crop in a given PLSS section and

what crops are present in a PLSS section according to what has been classified from Landsat images.

#### *Strengths*

Strengths include using NDVI values as the basis for image classification, representing the most widely used vegetation index and highly correlated with photosynthetic activity relevant to our interest in classifying agricultural land uses (Pettorelli et al., 2005; USGS, 2011). The Landsat method uses an improved MLC and per-field classification approach compared to a previous study implementing a minimum distance method representing training data crop fields with a single NDVI pixel value at the label point or that was seemingly representative of the polygon that failed to account for variability within and across crop fields (Maxwell, Airola, et al., 2010; Maxwell, Meliker, et al., 2010). The sample a priori Bayesian weighting logic in MLC incorporates information regarding particular land use classes (e.g., cotton) having consistently dominated the agricultural landscape in Kern County throughout the study time period (Campbell & Wynne, 2011; USDA, 2014). The Landsat method also includes per-field classification integrating raster and vector data that accounts for within-field spectral variability as each pixel's captured spectral signature may be impacted by soil moisture, pests, and disease (De Wit & Clevers, 2004; Guerschman et al., 2003; Turker & Ozdarici, 2011). Through implementing a majority rule in per-field classification, the spatial autocorrelation of agricultural crop fields was addressed, as pixels close in proximity likely belong to the same agricultural crop field, essentially averaging out the noise caused by the typical salt-and-pepper effects of per-pixel MLC classifiers (Lu & Weng, 2007). Furthermore, (crop) fields were created using a local behavior-based image segmentation procedure, which implemented a watershed delineation method of merging/growing pixels across input bands (monthly NDVI images) exhibiting minimal variance, or spectrally similar pixels that are likely of the same land use (Yu et al., 2006). Although the vector fields used in per-field classification are typically parcels dividing the landscape (i.e., the segments used may not parallel crop field boundaries) (Hill, 1999; Lu & Weng, 2007), classified segments were dissolved according to crop type and PLSS section, which was meaningful in terms of linking the PUR database to agricultural crop fields. Sensitivity analyses of the accuracy assessment using different segment sizes also demonstrated similar results (phenological group kappas ranging from 0.70 to 0.75).

#### *Limitations and future research*

Misclassification of crops exhibiting similar phenological patterns was evident in some high omission and commission errors and potentially contributed to a higher proportion of LUS tier 1 matches in 1985 compared to Landsat among some crops. A potential contributing factor to phenological misclassification is the absence of November and December images within a given crop's growing season. A refined three-tier methodology incorporating the LUS and Landsat methods that harnesses the results regarding phenological misclassification and the differential capabilities of capturing temporary vs. permanent crops should be explored. Another limitation is associated with the hard classification used to assign each segment to one land use class. Although precautions were taken when creating training data (e.g., single-use LUS), the classification method could not account for intra-annual crop rotation, for example, when two or more crops are successively grown on the same field each year (Gliessman, 1985). Focusing the analysis on acquiring imagery to classify specific crop types with known planting and harvesting dates may minimize this issue.

More focused GIS exposure modeling of environmental health risks would improve exposure estimation by including information regarding how pesticides are transported through the air, soil, and water (Lam, 2012). Although a 500 m buffer has been used in previous studies for its relevance to pesticide drift, it could be replaced with a boundary (defined by a pesticide exposure threshold) delineated using a dispersion model to determine where applied pesticides are transported.

The utility of using NDVI land use signatures from one year to train images from another year is affected by spatiotemporal changes in NDVI values. Although we used a time series of NDVI images for training and classification, NDVI values are affected by remote sensing system characteristics, meteorological conditions, ecosystem disturbances, and seasonality (Forkel et al., 2013). NDVI values can be unreliable in sparsely vegetated areas due to soil reflectance (Pettoirelli et al., 2005). It would be valuable to explore other vegetation indices, e.g., Soil Adjusted Vegetation Index (SAVI) that adjusts for soil spectral variability (Sonnenschein, Kuemmerle, Udelhoven, Stellmes, & Hostert, 2011), and non-reliance on a single vegetation index for classification, but rather implementing sophisticated classification techniques that incorporate ancillary information (e.g., slope) to further enhance classification accuracy (Yu et al., 2006).

## Conclusions

The Landsat method is an improved GIS and remote sensing method that can be used to estimate individual-level agricultural pesticide exposure at geocoded locations. The Landsat method addresses the dynamic nature of agriculture, especially crop rotations, through providing an approach to match pesticide applications to crop fields classified from temporally concurrent Landsat images rather than outdated LUS crop fields from the standard GIS method in California. In a demonstration of the Landsat method's ability to improve pesticide exposure estimation using 1985 data in California, significantly more pesticide applications, particularly among temporary crops, were matched to Landsat crops compared to LUS crops. This improvement is attributed to matching 1985 pesticide data to crops classified from 1985 Landsat images as opposed to 1990 LUS crops, which would be used in practice as it is the LUS conducted closest in time to the 1985 pesticide data. Future research should explore using the Landsat method as an exposure metric in an epidemiologic study as it is able to reconstruct historical exposure to specific chemicals at geocoded locations, and using ancillary data in Landsat image classification to improve land use classification and pesticide exposure estimation accuracy.

## Acknowledgments

The authors would like to acknowledge the following for providing publicly available data used for this study: the USGS and NASA for providing Landsat images, the California Department of Pesticide Regulation for providing Pesticide Use Reports, the California Department of Water Resources for providing land use surveys, the National Agriculture Imagery Program and Cal-Atlas Geospatial Clearinghouse for providing compressed county mosaic (CCM) imagery, and the U.S. Census Bureau for providing TIGER/Line© files.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.apgeog.2015.04.009>.

## References

- AgroPages. (2014). Agrodata crop protection database. <http://www.agropages.com/AgroData/>. Last accessed 01.08.13.
- Alavanja, M. C., Hoppin, J. A., & Kamel, F. (2004). Health effects of chronic pesticide exposure: cancer and neurotoxicity. *Annual Review of Public Health*, 25, 155–197. <http://dx.doi.org/10.1146/annurev.publhealth.25.101802.123020>.
- Blair, A., Ritz, B., Wesseling, C., & Beane Freeman, L. (2015). Pesticides and human health. *Occupational and Environmental Medicine*, 72(2), 81–82. <http://dx.doi.org/10.1136/oemed-2014-102454>.
- Campbell, J. B., & Wynne, R. H. (2011). *Introduction to remote sensing* (5th ed.). New York, NY: The Guilford Press.
- CDPR (California Department of Pesticide Regulation). (2014). Pesticide use reporting (PUR). <http://www.cdpr.ca.gov/docs/pur/purmain.htm>. Last accessed 01.08.13.
- CDWR (California Department of Water Resources). (2009). Metadata for the 2009 Eastern Fresno County land use survey data. <http://www.water.ca.gov/landwateruse/docs/landusedata/shapes/09eFR.zip>. Last accessed 01.08.13.
- CDWR (California Department of Water Resources). (2014). Land use survey overview. <http://www.water.ca.gov/landwateruse/lusrvymain.cfm>. Last accessed 01.08.13.
- Centre for Ecology and Hydrology. (2015). Land cover mapping. <http://www.ceh.ac.uk/landcovermapping.html>. Last accessed 01.04.15.
- Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5), 893–903. <http://dx.doi.org/10.1016/j.rse.2009.01.007>.
- Chavez, P. S. (1996). Image-based atmospheric corrections-revisited and improved. *Photogrammetric Engineering and Remote Sensing*, 62(9), 1025–1035.
- Chen, X., Li, B., & Allen, M. F. (2010). Characterizing urbanization, and agricultural and conservation land-use change in Riverside County, California, USA. *Annals of the New York Academy of Sciences*, 1195(s1), E164–E176. <http://dx.doi.org/10.1111/j.1749-6632.2009.05403.x>.
- Clark Labs. (2014). *IDRISI Selva*. Worcester, Massachusetts, United States.
- De Wit, A. J. W., & Clevers, J. G. P. W. (2004). Efficiency and accuracy of per-field classification for operational crop mapping. *International Journal of Remote Sensing*, 25(20), 4091–4112. <http://dx.doi.org/10.1080/01431160310001619580>.
- Dich, J., Zahm, S. H., Hanberg, A., & Adami, H. O. (1997). Pesticides and cancer. *Cancer Causes Control*, 8(3), 420–443. <http://dx.doi.org/10.1023/A:1018413522959>.
- Esri. (2014). *ArcGIS 10.1*. Redlands, California, United States.
- Fera. (2015). Pesticide usage surveys. <https://secure.fera.defra.gov.uk/pusstats/surveys/index.cfm>. Last accessed 01.04.15.
- Food and Agriculture Organization. (2013). FAOSTAT glossary. <http://faostat.fao.org/site/375/default.aspx>. Last accessed 01.03.14.
- Forkel, M., Carvalhais, N., Verbesselt, J., Mahecha, M. D., Neigh, C. S. R., & Reichstein, M. (2013). Trend change detection in NDVI time series: effects of inter-annual variability and methodology. *Remote Sensing*, 5(5), 2113–2144. <http://dx.doi.org/10.3390/rs5052113>.
- Franklin, C., & Worgan, J. (2005). *Occupational and residential exposure assessment for pesticides*. Hoboken, NJ: Wiley.
- Gliessman, S. R. (1985). Multiple cropping systems: a basis for developing an alternative agriculture. In Office of Technology Assessment (Ed.), *Innovative technologies for lesser developed countries* (pp. 69–86). U.S. Congress, Washington D.C.
- Godfrey, L., Rosenheim, J., & Goodell, P. (2000). Cotton aphid emerges as major pest in SJV cotton. *California Agriculture*, 54(6), 26–29. <http://dx.doi.org/10.3733/ca.v054n06p26>.
- Greene, S. A., & Pohanish, R. P. (2005). *Sittig's handbook of pesticides and agricultural chemicals*. Norwich, N.Y: William Andrew Publishing.
- Guerschman, J. P., Paruelo, J. M., Bella, C. D. I., Giallorenzi, M. C., & Pacin, F. (2003). Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. *International Journal of Remote Sensing*, 24(17), 3381–3402. <http://dx.doi.org/10.1080/0143116021000021288>.
- Gunier, R. B., Harnly, M. E., Reynolds, P., Hertz, A., & Von Behren, J. (2001). Agricultural pesticide use in California: pesticide prioritization, use densities, and population distributions for a childhood cancer study. *Environmental Health Perspectives*, 109(10), 1071–1078. PMID: 11689348.
- Gunier, R. B., Ward, M. H., Airola, M., Bell, E. M., Colt, J., Nishioka, M., et al. (2011). Determinants of agricultural pesticide concentrations in carpet dust. *Environmental Health Perspectives*, 119(7), 970–976. PMID: 21330232.
- Harrison, J. L. (2006). 'Accidents' and invisibilities: scaled discourse and the naturalization of regulatory neglect in California's pesticide drift conflict. *Political Geography*, 25(5), 506–529. <http://dx.doi.org/10.1016/j.polgeo.2006.02.003>.
- Hill, R. A. (1999). Image segmentation for humid tropical forest classification in Landsat TM data. *International Journal of Remote Sensing*, 20(5), 1039–1044. <http://dx.doi.org/10.1080/014311699213082>.
- Kern County Assessor. (2012). Use codes. <http://recorder.co.kern.ca.us/assessor/FYI.php?page=usecodes>. Last accessed 01.04.13.
- Lam, N. S. N. (2012). Geospatial methods for reducing uncertainties in environmental health risk assessment: challenges and opportunities. *Annals of the Association of American Geographers*, 102(5), 942–950. <http://dx.doi.org/10.1080/00045608.2012.674900>.

- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174. PMID: 843571.
- Liebman, M., & Dyck, E. (1993). Crop rotation and intercropping strategies for weed management. *Ecological Applications*, 92–122. <http://dx.doi.org/10.2307/1941795>.
- Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5), 823–870. <http://dx.doi.org/10.1080/01431160600746456>.
- Martinuzzi, S., Gould, W. A., & Ramos González, O. M. (2007). Creating cloud-free Landsat ETM+ data sets in tropical landscapes: cloud and cloud-shadow removal. <http://www.fs.fed.us/global/iitf/pubs/iitf-gr32.pdf>. Last accessed 01.08.13.
- Maxwell, S. K. (2011). Downscaling pesticide use data to the crop field level in California using Landsat satellite imagery: Paraquat case study. *Remote Sensing*, 3(9), 1805–1816. <http://dx.doi.org/10.3390/rs3091805>.
- Maxwell, S. K., Airola, M., & Nuckols, J. R. (2010). Using Landsat satellite data to support pesticide exposure assessment in California. *International Journal of Health Geographics*, 9(46). <http://dx.doi.org/10.1186/1476-072X-9-46>.
- Maxwell, S. K., Meliker, J. R., & Goovaerts, P. (2010). Use of land surface remotely sensed satellite and airborne data for environmental exposure assessment in cancer research. *Journal of Exposure Science and Environmental Epidemiology*, 20(2), 176–185. <http://dx.doi.org/10.1038/jes.2009.7>.
- Mueller-Warrant, G. W., Whittaker, G. W., Griffith, S. M., Banowetz, G. M., Dugger, B. D., Garcia, T. S., et al. (2011). Remote sensing classification of grass seed cropping practices in Western Oregon. *International Journal of Remote Sensing*, 32(9), 2451–2480. <http://dx.doi.org/10.1080/01431161003698351>.
- National Atlas. (2014). The public land survey system (PLSS). Last accessed 01.04.13 [http://www.nationalatlas.gov/articles/boundaries/a\\_plss.html](http://www.nationalatlas.gov/articles/boundaries/a_plss.html).
- Natural Resources Geographic Information System Library. (2015). <https://programs.iowadnr.gov/nrgislibx/>. Last accessed 01.04.15.
- Nuckols, J. R., Gunier, R. B., Riggs, P., Miller, R., Reynolds, P., & Ward, M. H. (2007). Linkage of the California pesticide use reporting database with spatial land use data for exposure assessment. *Environmental Health Perspectives*, 115(5), 684–689. PMID: 17520053.
- Pesticide Action Network. (2013). PAN pesticides database – California pesticide use. [http://www.pesticideinfo.org/Search\\_Use.jsp](http://www.pesticideinfo.org/Search_Use.jsp). Last accessed 01.08.13.
- Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, 20(9), 503–510. <http://dx.doi.org/10.1016/j.tree.2005.11.006>.
- Phillips, J. D. (1988). Nonpoint source pollution and spatial aspects of risk assessment. *Annals of the Association of American Geographers*, 78(4), 611–623. <http://dx.doi.org/10.1111/j.1467-8306.1988.tb00234.x>.
- Richards, J. A. (2013). *Remote sensing digital image analysis: An introduction* (5th ed.). New York: Springer.
- Rull, R. P., Gunier, R., Von Behren, J., Hertz, A., Crouse, V., Buffler, P. A., et al. (2009). Residential proximity to agricultural pesticide applications and childhood acute lymphoblastic leukemia. *Environmental Research*, 109(7), 891–899. <http://dx.doi.org/10.1016/j.envres.2009.07.014>.
- Rull, R. P., & Ritz, B. (2003). Historical pesticide exposure in California using pesticide use reports and land-use surveys: an assessment of misclassification error and bias. *Environmental Health Perspectives*, 111(13), 1582–1589. PMID: 14527836.
- Rull, R. P., Ritz, B., & Shaw, G. M. (2006a). Neural tube defects and maternal residential proximity to agricultural pesticide applications. *American journal of epidemiology*, 163(8), 743–753. <http://dx.doi.org/10.1093/aje/kwj101>.
- Rull, R. P., Ritz, B., & Shaw, G. M. (2006b). Validation of self-reported proximity to agricultural crops in a case-control study of neural tube defects. *Journal of Exposure Science and Environmental Epidemiology*, 16(2), 147–155. <http://dx.doi.org/10.1038/sj.jea.7500444>.
- Simoniello, T., Carone, M. T., Lanfredi, M., Macchiato, M., & Cuomo, V. (2004). Landscape-scale characterization of vegetation phenology using AVHRR-NDVI and Landsat-TM data. In *Remote sensing* (pp. 644–651). International Society for Optics and Photonics.
- Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., & Macomber, S. A. (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote sensing of environment*, 75(2), 230–244. [http://dx.doi.org/10.1016/S0034-4257\(00\)00169-3](http://dx.doi.org/10.1016/S0034-4257(00)00169-3).
- Sonnenschein, R., Kuemmerle, T., Udelhoven, T., Stellmes, M., & Hostert, P. (2011). Differences in Landsat-based trend analyses in drylands due to the choice of vegetation estimate. *Remote sensing of environment*, 115(6), 1408–1420. <http://dx.doi.org/10.1016/j.rse.2011.01.021>.
- Sullivan, P. (2003). *Intercropping principles and production practices*. <https://attra.ncat.org/publication.html>. Last accessed 01.08.13.
- Suzuki, R., Nomaki, T., & Yasunari, T. (2001). Spatial distribution and its seasonality of satellite-derived vegetation index (NDVI) and climate in Siberia. *International Journal of Climatology*, 21(11), 1321–1335. <http://dx.doi.org/10.1002/joc.653>.
- Turker, M., & Ozdarici, A. (2011). Field-based crop classification using SPOT4, SPOT5, IKONOS and QuickBird imagery for agricultural areas: a comparison study. *International Journal of Remote Sensing*, 32(24), 9735–9768. <http://dx.doi.org/10.1080/01431161.2011.576710>.
- U.S. Census Bureau. (2014). TIGER products. <http://www.census.gov/geo/maps-data/data/tiger.html>. Last accessed 01.01.14.
- University of California Statewide Integrated Pest Management Program. (2006). *Pest management guidelines: Alfalfa crop rotation*. <http://www.ipm.ucdavis.edu/PMG/r1900811.html>. Last accessed 01.01.13.
- USDA (U.S. Department of Agriculture). (2003). *USDA Forest Service Gen. Tech. Rep. PSW-GTR-187: The Central Valley Region*. [http://www.fs.fed.us/psw/publications/documents/psw\\_gtr187/psw\\_gtr187\\_05.pdf](http://www.fs.fed.us/psw/publications/documents/psw_gtr187/psw_gtr187_05.pdf). Last accessed 01.01.14.
- USDA (U.S. Department of Agriculture). (2013). Soil tillage and crop rotation. <http://www.ers.usda.gov/topics/farm-practices-management/crop-livestock-practices/soil-tillage-and-crop-rotation.aspx>. Last accessed 01.01.14.
- USDA (U.S. Department of Agriculture). (2014). Census of agriculture. <http://www.agcensus.usda.gov/index.php>. Last accessed 01.01.14.
- USDA Farm Service Agency Aerial Photography Field Office. (2012). Kern county compressed county mosaic (CCM). In *National agriculture imagery program (NAIP) via Cal-Atlas geospatial clearinghouse*. Salt Lake City, Utah.
- USGS (U.S. Geological Survey). (2011). NDVI, the foundation for remote sensing phenology. [http://phenology.cr.usgs.gov/ndvi\\_foundation.php](http://phenology.cr.usgs.gov/ndvi_foundation.php). Last accessed 01.04.13.
- USGS (U.S. Geological Survey). (2013). *Landsat path/row shapefiles*. [http://landsat.usgs.gov/tools\\_wrs-2\\_shapefile.php](http://landsat.usgs.gov/tools_wrs-2_shapefile.php). Last accessed 01.04.13.
- USGS (U.S. Geological Survey). (2014). *Landsat: A global land-imaging mission*. <http://landsat.usgs.gov/>. Last accessed 01.02.14.
- Van Niel, T. G., & McVicar, T. R. (2004). Determining temporal windows for crop discrimination with remote sensing: a case study in south-eastern Australia. *Computers and Electronics in Agriculture*, 45(1), 91–108. <http://dx.doi.org/10.1016/j.compag.2004.06.003>.
- Vassiliou, A., Boulianne, M., & Blais, J. (1988). On the application of averaging median filters in remote sensing. *Geoscience and Remote Sensing, IEEE Transactions on*, 26(6), 832–838. <http://dx.doi.org/10.1109/36.7714>.
- Wan, N. (2015). Pesticides exposure modeling based on GIS and remote sensing land use data. *Applied Geography*, 56, 99–106. <http://dx.doi.org/10.1016/j.apgeog.2014.11.012>.
- Ward, M. H., Nuckols, J. R., Weigel, S. J., Maxwell, S. K., Cantor, K. P., & Miller, R. S. (2000). Identifying populations potentially exposed to agricultural pesticides using remote sensing and a geographic information system. *Environmental Health Perspectives*, 108(1), 5–12. PMID: 10622770.
- Wood, A. (2010). *Alan Wood compendium of pesticide common names*. [http://www.alanwood.net/pesticides/class\\_pesticides.html](http://www.alanwood.net/pesticides/class_pesticides.html). Last accessed 01.08.13.
- Yale Center for Earth Observation. (2013). *Yale guide to landsat 8 image processing*. <http://www.yale.edu/ceo/Documentation/Landsat%208%20image%20processing.pdf>. Last accessed 01.08.13.
- Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., & Schirokauer, D. (2006). Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering and Remote Sensing*, 72(7), 799 <http://dx.doi.org/10.14358/PERS.72.7.799>.