REMOTE SENSING & GIS APPLICATIONS FOR DRAINAGE DETECTION AND MODELING IN AGRICULTURAL WATERSHEDS

Samapriya Roy

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Science
in the Department of Earth Sciences,
Indiana University

July 2013

Accepted by the Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

Lin Li, Ph.D., Chair

Master's Thesis Committee

Edward Randall Bayless, Ph.D.

Pierre-Andre Jacinthe, Ph.D.

DEDICATION

I would like to dedicate this work to my family and friends, those who have in some sense have allowed me to think freely and act on it. It is often in their company that I yield to day dreams and convert them into ideas.

"Through the thickest times and thinnest hope,

We keep the best of dreams in our mind always afloat.

I learn to tread silently, among weary pathways in my daydreams,

Forever more I come as a student to you, my mentor, and a value to keep."

ACKNOWLEDGEMENTS

It is difficult to get second chances to prove something; I think first and foremost I would like to thank all those who believed in me to lend me that faith. I would like to thank my family and my friends, in some way or the other they have taught me to keep your spirits high, even when things don't work out the way you want it to be.

I would like to take this opportunity and thank my graduate committee, especially my graduate chair, Dr. Lin Li, I believe I have learned more than just doing fine work with him, I have valued our conversations more than a lot of things and I hope it has helped me learn and grow in every dimension.

I would also like to thank Dr. Randall Bayless and Dr. Pierre Jacinthe, my committee members, my professors and my mentors at time, for being patient with me. I would like to extend my appreciation as they allowed me enough time to work with the problem and encouraged me to tackle the problem from different directions. I have always had the opportunity.

I would specially like to thank Dr. Nancy Baker from USGS for meeting me the first time I came up with this idea last summer and in instilling the need for something to be done in that direction. Your encouragement is highly appreciated and your role as an external mentor on the project has truly helped me in developing the project in different directions. I would also like to mention John Milbrun from Hancock County office and David S. Nail for the USGS for providing me with a variety of subsidiary and relevant datasets which were required for my work.

In the course of my work here I have learned to value critics, insights and helpful doubts from my friends and peers. This has helped me in moving in a positive direction and for gaining a set of extra eyes when I needed someone to go over my thesis sections.

My interactions with some of my professors have always been valuable to me which includes conversation with Dr. Gabriel Filippelli, Jennifer Nelson and others who have had many insightful conversations with me. Teaching my classes had always been more interesting as I learned from my valued mentors and professors and I have tried to imbibe some of these skills in my own classes.

I would also like to thank the Department Chair, Dr. Kevin Mandernack for adding that extra bit of guidance I needed at time. I earnestly thank faculty, staff and students here in the Department of Earth sciences for helping me with all that was required and for everything that kept me going. I appreciate the financial support from the Department of Earth Sciences which has helped me through my time here.

ABSTRACT

Samapriya Roy

REMOTE SENSING & GIS APPLICATIONS FOR DRAINAGE DETECTION AND MODELING IN AGRICULTURAL WATERSHEDS

The primary objective of this research involves mapping out and validating the existence of sub surface drainage tiles in a given cropland using Remote Sensing and GIS methodologies. The process is dependent on soil edge differentiation found in lighter versus darker IR reflectance values from tiled vs. untiled soils patches. Data is collected from various sources and a primary classifier is created using secondary field variables such as soil type, topography and land Use and land cover (LULC). The classifier mask reduces computational time and allows application of various filtering algorithms for detection of edges.

The filtered image allows an efficient feature recognition platform allowing the tile drains to be better identified. User defined methods and natural vision based methodologies are also developed or adopted as novel techniques for edge detection. The generated results are validated with field data sets which were established using Ground Penetration Radar (GPR) studies. Overlay efficiency is calculated for each methodology along with omission and commission errors. This comparison yields adaptable and efficient edge detection techniques which can be used for similar areas allowing further development of the tile detection process.

Lin Li, PhD., Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADS..Agricultural Drainage Systems

USDA...United States Department of Agriculture

USGS..United States Geological Survey

NAIP..National Agriculture Initiative Program

NASS...National Agriculture Survey Statistic

GPR...Ground Penetrating Radar

GIS...Geographic Information System

IR...Infrared

CIR...Color Infrared

HUC...Hydrologic Unit Code

DOQQ...Digital Ortho Quarter Quads

LULC..Land Use and Land Cover

NDVI...Normalized Difference Vegetative Index

DEM..Digital Elevation model

SSURGO...Soil Survey Geographic Database

CTI...Compound Topographic Index

TRI..Terrain Ruggedness Index

PCD..Phase Congruency Detection

CORF...Combination of Receptive fields

INTRODUCTION

Agricultural watersheds are often drained by agricultural drainage systems (ADS) which are designed to allow easy transport of water from land and benefit agriculture and agricultural yield due to the reduced level of water in the fields (Oosterbaan, 1994). These ADS can be surface drains and ditches which allow the water to pass on the surface and are called "surface drainage systems", or tile drains which allow the water to pass underneath the soil and are called "subsurface drainage systems". Both surface and subsurface systems drain into an internal low, the overall water availability in the field or area of interest, and into an external or main drain transport which transports the water to the outlet or riverine system. Tile drainage has increased the ratio of subsurface to surface drainage creating major concern about the hydrologic consequences caused by these subsurface agricultural drainage systems and thus subsurface tile drain systems are the focus of the current study. The location information of these tile drain systems are limited and the tile drain placement layouts are complex, making detection of subsurface systems a valuable tool in assessment of field hydrology and in understanding flow of water through these agricultural areas. This further improves understanding of flow routing for control and assessment of water quality and quantity parameters in a given agricultural watershed.

Past investigations into location of these tile drain systems have yielded limitations in terms of a singular approach and the wide applicability in terms of designed methodology (Allred et al., 2004; Verma et al., 1996; Varner et al., 2003; Naz and Bowling 2008). However owing to the extent of the problem of detection and the area under consideration, remote sensing has been established as a key strategy. It plays a valuable role in reducing human intervention and constraints while handling a large area and has thus been developed further for our current study.

Hydrological Impacts of Subsurface ADS

Tile drain systems are generally placed about 0.6 to 1.2 m depth underneath the soil surface and made from ceramic, clay or corrugated plastic materials. The function and location of these subsurface drainage systems vary with the inherent soil saturation and the amount of water drained by the systems. Subsurface drainage improves the moisture and aeration conditions of the soil and has been known to reduce the total surface runoff by increasing total infiltration (Stillman et al., 2006). Subsurface drainage systems were originally placed to reduce the duration of excess soil water in the root zone, facilitate crop production and increase productivity.

From 1997 to 2007 the average total agricultural productivity in United States increased tremendously. Particularly the corn production has increased from 8,732,478,098 bushels in 1997 to 12,738,519,330 bushels in 2007 (USDA Census of Agriculture 2007). Increased agricultural productivity has promoted large areas with saturated soil conditions to be converted into tiled agricultural systems (Table 1). Other benefits include reduced surface runoff and hence soil erosion, and improved water filtration in comparison to untiled areas (Skaggs and van Schilfgaarde, 1999).

Table 1. The number of wetlands converted to agricultural type (1980-88 to 2005)

(Indiana NWI Update Final Report, 2010)

	Agriculture	
	Acres	%
Fully Converted	25,023.05	79.20%
Partially Converted	7,895.24	57.12%
Total	32,918.29	**72.48%**

However, there are concerns about the watershed hydrology and water quality (Naz and Bowling, 2008). The interactions between groundwater, surface water and subsurface drainage systems clearly elucidate the importance of a proper simulation and flow routing and distribution model to understand water quantity and solute transfer dynamics (Rozemeijer et al., 2010).

Though these studies have included the amount of water removed by tile drains, the timing of flow through a tile drainage system has not been incorporated into the design criteria (Walter et al., 1979). The absence of a control mechanism from these tiled fields have been linked to large loads of residual nitrogen being exported to surrounding waterways, imposing a problem to the overall water quality of the system (Goswami et al., 2009).

Functional Assessment of Tile Drainage Systems

It is evident that assessing the functional and environmental benefits of a tile drainage system is largely dependent on understanding the spatial distribution and layout of underlying tile drainage networks. The first step toward this understanding is to examine the structure of a tile drainage system. Based on previous studies and the recommendations from the Indiana Drainage Handbook (1996), an optimal tile drain layout (Fig. 1) aims to provide an adequate and uniform drainage to a given field (Wright and Sands, 2001).

The layout patterns in earlier studies include parallel spaced drains with fixed spacing and minimum slope variation to targeted drainage (Wright and Sands, 2001). However such layout patterns and practices were not followed earlier and hence large sections of tile drain networks consist of irregular networks (Cooke et al., 2001).

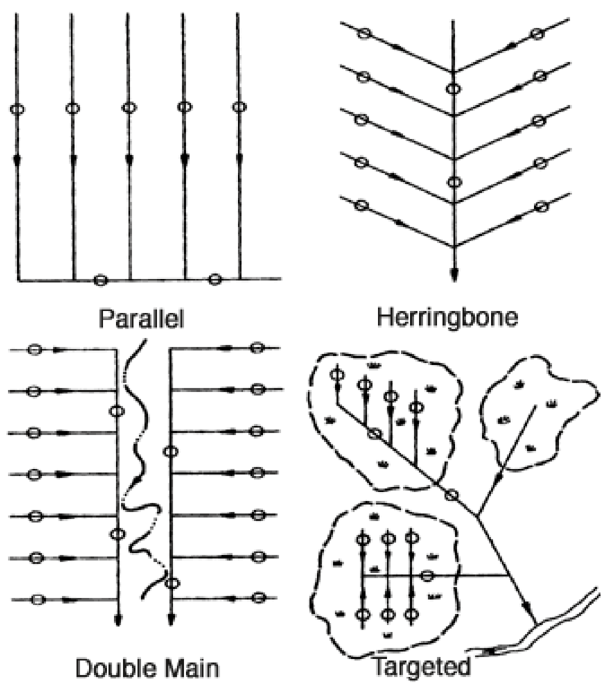

Fig. 1 Tile drain patterns in agricultural watersheds (adopted from Wright and Sands, 2001)

This led to differences in the actual tile intensity in the given area and the revised statistics for percentage area that has been tiled in the 1987 and 1992 agricultural census report (Table 2). For example, the USDA-ERS 1987 report stated that 70% of drained land in Indiana is subsurface tile drained, and the 1992 USDA agricultural census generated similar results (Table 2).

Table 2. Estimated percent cropland with subsurface drainage for Indiana (USDA 1992)

Indiana	
(A) Total cropland 1992 (ac)	13,370,000
(B) Drained cropland (50%of total cropland)	6,685,000
(C) Drained cropland with subsurface drainage (82%of total drained cropland)	5,481,700

This inconsistency makes understanding the flow dynamics and the role of subsurface tile drains in their response to agricultural flow viable. Additionally, the planning stages for new tile systems require information about the existing drains and their placement, spacing and tillage intensity including both partially functional and nonfunctional tiles. However such information on subsurface tiles is currently lacking. The identification of tile train patterns depends on a methodology to be developed and adapted.

Earlier studies on tile drain identification (Allred et al., 2004; Verma et al., 1996; Varner 2003; Naz and Bowling 2008) have included ideal tile drain layout, which are regularly placed and do not include irregular tile drain patterns. The current understanding of tile layout does not delve into the problem of larger existing irregular tile systems (Cooke et al., 2001). Another distinct classification that distinguishes tile drains systems based on their functionality, groups subsurface tiles into controlled and uncontrolled flow systems. The ideal situation under any given conditions would be a regularly placed controlled flow system which allows the farmer to control the time period and quantity of discharge (Purdue Climate Report, 2008). This has instigated interest in understanding existing tile systems for required modification and augmentation (Allred et al., 2004). The improved ADS will inculcate existing systems while optimizing tile density and spacing. However, most tile drain systems are truly irregular and have uncontrolled flow systems (Table 3).

Table 3. Punnett square representation of tile pattern and control

	Regular	Irregular
Controlled	Regular, Controlled	Irregular, Controlled
Uncontrolled	Regular, Uncontrolled	Irregular, Uncontrolled

The increase in agricultural practices will lead to substantial increase in tile intensity (Number of tiles/Unit Area) and the total tiled land area. This would require additional tiling operations for previously undrained areas or in areas with existing nonfunctioning and partially functioning drains which have to be replaced or augmented as previously discussed. The lack of existing tile structures coupled with the need for increase tiling provides added impetus for understanding the location of the tile drains.

STATUS OF TILE DRAIN DETECTION

Geophysical approaches

Exploration has been conducted with limited geophysical methods to determine the location and pattern of subsurface drainage networks (Allred et al., 2004). The research was conducted as grid surveys in areas of southwest, central, and northwest Ohio at eleven test plots containing these subsurface drainage systems and the detection efficiency was compared based on the methodology used (Allred et al., 2004). The method included geomagnetic surveying, electromagnetic induction, resistivity, and Ground Penetrating Radar (GPR) studies which were conducted for fields containing clay tile and corrugated plastic tubing drainage pipe. The average effectiveness was found to be 81 % in terms of locating the subsurface drainage pipe (Allred et al., 2004). Although the techniques were accurate, they were clearly inefficient considering the expanse of the problem, total time required and related aspects of how cost ineffective this method could be. In addition, the scale of the problem and the varying response types and system patterns pose limitations on the methodology.

Remote Sensing approaches

Remote Sensing (RS) and Geographic Information System (GIS) have allowed us to develop cost effective tools for the delineation of these tile drains. The process reduces the time required for tile drain detection, and a large area can be covered. Tile drain mapping using RS is based on the fact that the soil over efficiently draining tiles should dry faster than the soil at other locations in the field. This creates a high reflectance in the infrared (IR) region of the radiation spectrum. The IR range of the radiation spectrum is very sensitive to soil moisture, and these variations in near IR range (0.7-1.3 μm) and mid IR range (1.3 μm) reflectance can be captured and differentiated (Verma et al., 1996). Other soil properties affecting the reflectance such as soil texture, the roughness, and the presence of organic content allow for narrowing down the area of study. The strong relationship between soil texture and reflectance value has been established: coarse, sandy soils which are well drained tend to have higher reflectance values than poorly drained fine textured soils. Tile drains are primarily found in poorly drained fine textured soils. This creates the relative difference in reflectance, which could be used to study the locations of tile drain systems.

The choice of the wavelength range for remote sensing images depends on the response of moisture to light in particular wavelengths and the possibility of ascertaining drainage edges from these buried subsurface drainage patterns. Verma et al. (1996) used color infrared imagery (CIR) acquired for a larger time period of March to April 1984 at a spatial resolution of about 1m. This CIR based study suggested that the image used by Verma et al. (1996) had uniformly light gray tone and indicated drainage stresses more efficiently. Varner (2003) used RDACSII multispectral sensor data, and the RDACSH3 hyperspectral data with 120 bands from 471 to 828nm which was used to simulate IKONOS data. The multispectral and simulated data at a spatial resolution of 1m was suitable for discerning these drainage features, but limited in temporal scale owing to cloud conditions. The study by Naz and Bowling (2008) benefitted from high spatial resolution for pan chromatic images between 0.15-1m and aerial color (NAIP) imagery at the 1m resolution.

The accuracy and efficiency for detection of subsurface tiles relies on remote sensing methodologies to be used. Verma et al. (1996) primarily used band transformation and manual digitization for visually delineation of tile drain features. This method is no doubt time consuming. Varner (2003) developed four distinct methodologies. The first is contrast enhancement followed by Kernel based convolutions and filtering using 3 x 3 kernels. The second is a hybrid soil guided methodology involving identification of areas with hydric soils, clipping the raster layer based on the location of hydric soils and applying a 3 x 3 median filter for non-directional edge enhancement. This method benefits by limiting the area of application and thereby increasing computational efficiency of detection of these drainage patterns. The third method is a combination of PCA application coupled with edge enhancements using horizontal, vertical and diagonal filters which were then layer stacked to get better response for features of interest; and the PCA methodology allowed for a better result with over 77% to 81% accuracy. The last method is to utilize the panchromatic data for the delineation. The overall accuracy of these methods reached about 77% though application was severely limited by the time frame of the data collected. Among image processing and edge detection methods mentioned above, none of the methodologies used are truly automatic when it comes to detection and vectorization.

When dealing with regular tile systems, Naz and Bowling (2008) applied automatic delineation and vectorization to high spatial resolution panchromatic and natural color imagery, and resulted in an improved accuracy even with the absence of hyperspectral data used by earlier studies (Varner, 2003). In this study, automated convolution with varying Kernel sizes was used to convolve the image and edge detection. In order to address the displacements between pixels, the Hough transformation was used to extract regular features likes lines, circles and ellipses. The

Hough transformation is an edge detection algorithm capable of extracting feature boundaries (in this case linear features) without losing any continuity and was found to be computationally suitable for smaller scale.

Although different methods for detection of drainage tiles are available, it is not easy to conclusively determine the efficacy of one method over the other owing to the difference in the type of datasets and the time period of remote sensing data acquisition. Furthermore, multiple points of interest have been excluded from all previous studies, such as ignoring large scale irregular tile drain systems and a mixture of both functional and nonfunctional tile drain. The edge detection method could benefit in conjunction with high resolution digital elevation models (DEM). The earlier work cumulatively shapes the outline of the problem but must be augmented by better datasets and methods. The current study aims to carve a way for an effective, automated and adaptable tile drainage methodology.

OBJECTIVES

The overall theme of this research is

- To be able to identify and detect tile drainage pattern in subsurface drainage systems predominant in agricultural watersheds, using Remote Sensing and GIS. Specific objectives for the same project included the following.
- To compare and develop user defined techniques along with novel applications of image processing including user defined kernel operator, phase congruency detection (PCD) and combination of receptive fields (CORF) to tile detection problem.
- To examine the efficiency and accuracy of predicted models and comparison using visual comparison and intersect analysis; and to conduct ground validation data assessment and overlay for efficiency analysis of results.

METHODOLOGIES

Study Site

The study site is located in the Hancock County, Indiana. The county has a total area of about 195,200 acres, 686 farms with the total land area being 171,673 acres (USDA Agricultural Census, 2007). The location and choice of this county for our study is attributed to the fact that the county is primarily agricultural and main producers of corn, soybeans and wheat. The study site is a HUC10 (0512020404) watershed called the Little Sugar Creek-Sugar Creek Watershed situated within the Driftwood Watershed which has an approximate area of about 84751.45 acres lying within the Hancock county and covering about 43% of the total area. The chosen area has a distributed land use and land cover (Fig. 2).

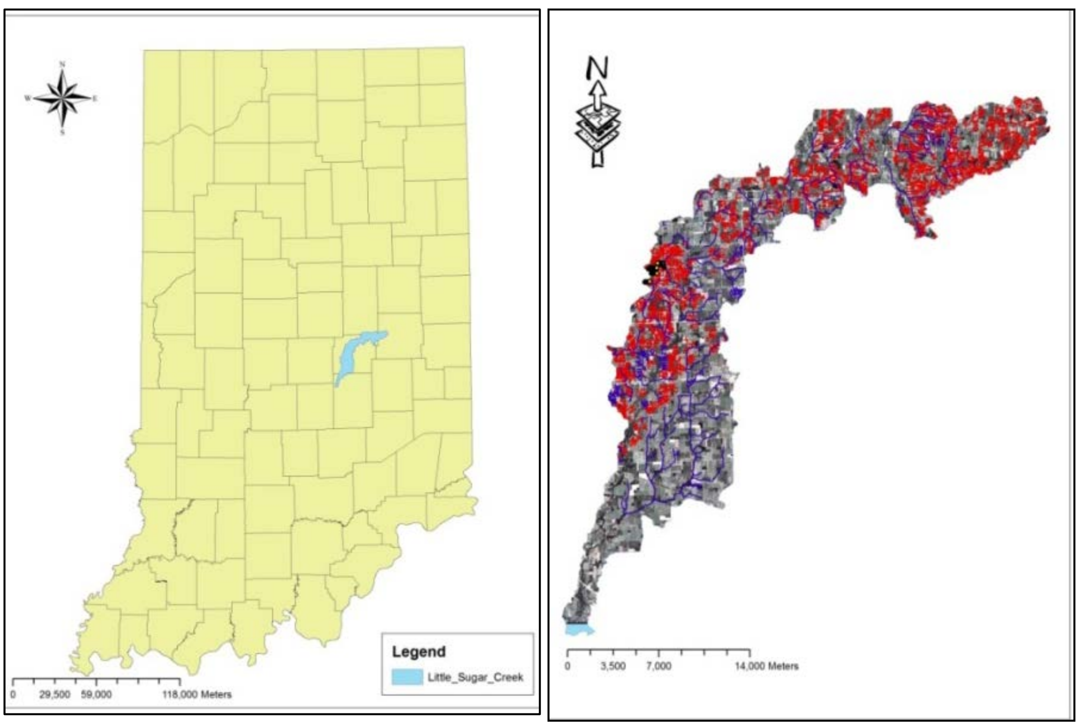

Fig. 2 Left: Area of interest Right: Sugar Creek (0512020404)

Previous studies were conducted in the study area by USGS for detection of tile drains using remote sensing (Fig. 3), and also field verification studies using GPR for validation of tile outlets and their point of intersection with regulated tiles (Fig. 4).

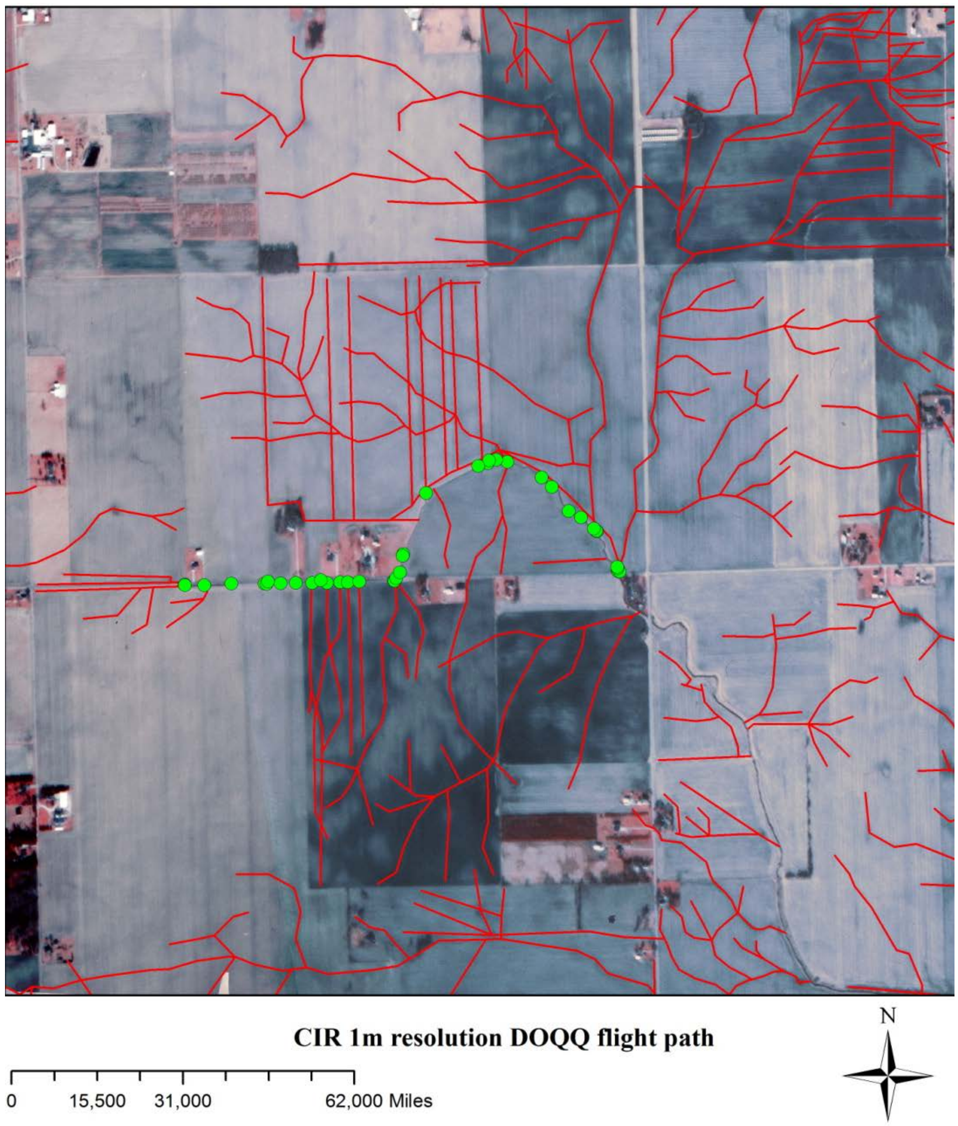

Fig. 3 Quad tile for Area of interest (CIR 1m resolution DOQQ flight path) overlaid with delineated tile drain

The soil profile for the county dictates the functionality of putting the tiles since the placement of tile systems are based on soil saturation and the benefit that can be achieved from subsurface drainage systems. The soils in the county have been classified based on their drainage classes for easy understanding of their function and adaptability to agricultural practices. For somewhat poorly drained soil, water is removed slowly so that the soil is wet at a shallow depth for significant periods during the growing season. Wetness negatively affects the growth of crop plants unless drainage is provided. These soils often have a slowly permeable limiting layer; for poorly drained soil, water is removed so slowly that the soil is wet at shallow depths periodically during the growing season or remains wet for long periods. Most crop plants cannot be grown unless artificial drainage is provided; for very poorly drained soil, water is removed so slowly that free water remains at or above the ground surface during much of the growing season. Most crop plants cannot be grown unless artificial drainage is provided.

Based on the assessment of the soil conditions recommendations have been made about the drainage characteristics and the improvement from subsurface drainage in these poorly drained soil categories. The poorly drained soils in the Hancock county include the following

- Crosby Brookston association

  Deep somewhat poorly drained and very poorly drained formed in glacial till and hence impermeable. These is a majority percentage of the soil distribution (72.4% of the county)

- Miami Crosby Association

  Deep well drained and somewhat poorly drained, nearly level to strongly sloping silt loams and clay loams that formed in glacial till, on uplands (16.9% of the county).

Remote Sensing Datasets

For our current work, raster datasets to be used include SPOT images and color infrared imagery(CIR) from both NAIP and DOQQ sources. CIR images were used because of their high spatial and temporal resolutions..

SPOT 5

The SPOT 5 satellite consists of a high resolution geometrical instrument (HRG) which offers panchromatic imagery anywhere from about 2.5-5m and the multispectral mode in 10-20 m.

Table 4. SPOT 5 Panchromatic & Multispectral Image characteristics and resolution

Dataset	Band	Wavelength(µm)	Spatial Resolution(m)
SPOT 5	Pan	0.48-0.71	5
	Green	0.50-0.59	10
	Red	0.61-0.68	
	Near IR	0.78-0.89	
	Shortwave IR	1.58-1.75	20

SPOT 5 images are highly suitable for tile drain features detection owing to the high resolution of its panchromatic band, and the path overlay between LANDSAT derived products and SPOT tiles. SPOT 5 data can be used as a standardized imagery with substantial amount of data for larger delineation features to be clearly identified. These images were obtained from the state USGS office and were georeferenced to the NAD 83 Datum.

2005 Indiana Map Color Infrared Orthophotography

This dataset is consistent of near infrared (NIR) data in conjunction to 3 band RGB bands, and has been obtained for the same Digital Ortho Quarter Quads (DOQQ) tiles at 1m high resolution for the State of Indiana during March and April leaf-off conditions. This dataset was collected by the Indiana Geographic Information Council (IGIC) and projected to State Plane Indiana East and West, NAD83 and provided in MrSid compression format.

Table 5. NAIP Color Infrared Imagery characteristics and resolution

Dataset	Band	Wavelength(μm)	Spatial Resolution(m)
NAIP(CIR)	Blue	0.45-0.52	1 m
	Green	0.52-0.60	
	Red	0.60-0.69	
	Near Infrared	0.75-0.90	

Color infrared orthophotography was taken during leaf off conditions to eliminate most of the crop cover and achieve better delineation of sub surface drainage features.

Digital Ortho Quarter Quads (DOQQ)

Aerial Digital Ortho Quarter Quads (DOQQ) was collected for the study area in tiles using an infrared sensor and the bands could be used for delineation of allied features.

Table 6. DOQQ Imagery characteristics and resolution (Source: USGS)

Platform	Band	Wavelength(μm)	Spatial Resolution(m)
DOQQ Flight	Band 2	0.52 - 0.60	1-2 m
	Band 3	0.63 - 0.69	
	Band 4	0.76 - 0.90	

The benefit for this dataset includes the fact that these flights were scheduled post a rainfall event and the high spatial resolution which makes it perfectly suitable for our study (Verma, 1996; Varner, 2003). The high resolution imagery also allows for better delineation as already expected.

Ground Penetration Radar(GPR) Data

Ground penetration radar applies radio frequency electromagnetic waves for subsurface surveying where the reflectance and return from subsurface layers are received and interpreted for gathering information. GPR data were collected from April 5 to 8th 2004 as part of the National Water Quality Assessment Program (NAWQA) study to verify the locations of drainage tiles identified from aerial infrared red photographs.

Both 100 and 250 Megahertz (MHz) shielded, monostatic Mala GPR antenna were used for the study, and data were collected along eleven transects across corn and soybean fields. The data for low pass and high pass were collected and the linear transects were then utilized to gather enough information regarding the delineated tile drains for comparison. The GPR data are beneficial in flatter lands with small slope undulations and are valuable for identifying the edges and edge junctions for the identified tile drains. These datasets were obtained from the USGS Indiana Water Science Center and applied for our area of interest.

Composite DEM

The composite DEM plays an important role in the analysis, allowing us to perform hydrological analysis and to understand flow direction and flow accumulation patterns in these regions and deriving the soil indices. The natural flow accumulation should intersect with areas of maximum tiling and thus the DEM derivatives can also be used for measuring minute topological and soil morphological variations owing to the higher resolution and owing to relatively flatter terrain of the study area which limits the use of a coarser resolution DEM. The composite DEM was derived from a 10 m resolution DEM data and the point mesh data from high resolution LIDAR to create DEM at a spatial resolution of 3ft. Indiana has a fairly flat morphology making most coarse resolution DEM not able to capture the minute soil morphological features. Also a higher resolution

DEM can be utilized to better decision classifier using the flow models. This composite DEM was obtained from the County Office at Hancock County.

LULC DATA (Source: National Agricultural Statistics Service, [NASS])

Based on the soil conditions and land use, land classification survey was collected. The information on land use and approximate coverage of agricultural area allows for understanding the distribution of distribution of crop type and yield. This cultivated crop mask data layer has the 30 m spatial resolution, and covers the continental United States. It is based on cropland data layers from 2007 through 2011 with plans to update this crop mask data layer annually.

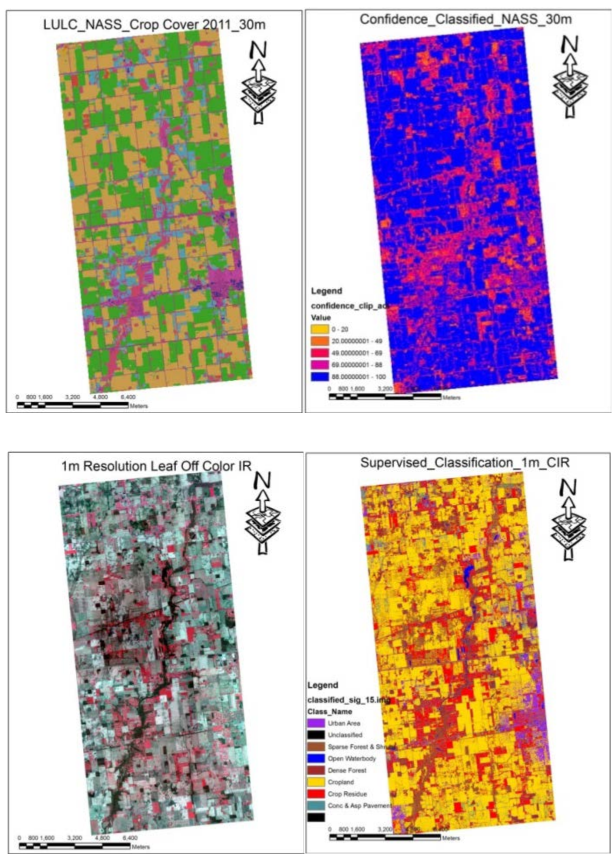

Fig. 4 30m Resolution NASS LULC and Confidence imagery and 1m CIR and classified LULC product

The derived LULC dataset allows us to isolate the built up area and other areas which are not of primary interest and feeds into the primary classifier system. The CIR derived product is utilized as a reference for comparison with the Landsat derived LULC data.

Vector Datasets

Vector data allow us to understand grid analysis and to overlay features which can be utilized for secondary analysis and application. The vector datasets were collected and new datasets were to be generated from the analysis for feature segmentation and detection (Table 7). Both raster and vector data sets were georeferenced to maintain consistency in the analyzed data and the processing which would follows.

Table 7. Vector Layer with attribute Data for secondary feature analysis

Source	Vector Layer	Attribute Data
HANCOCK Co. Office	Regulated Tile Drain	Tile Material
		Tile Length
		Tile Diameter
USGS	Indiana Wetland Inventory	Wetland Area
	HUC10 Watershed Outline	Drainage Area
	Census County Map	County Area & Location
	Baker Tile Grid	Tile Length
	GPR Point Validation	Point Location
	GPR Outlet validation	Point Location

Synergistic Analysis of Datasets

Overview

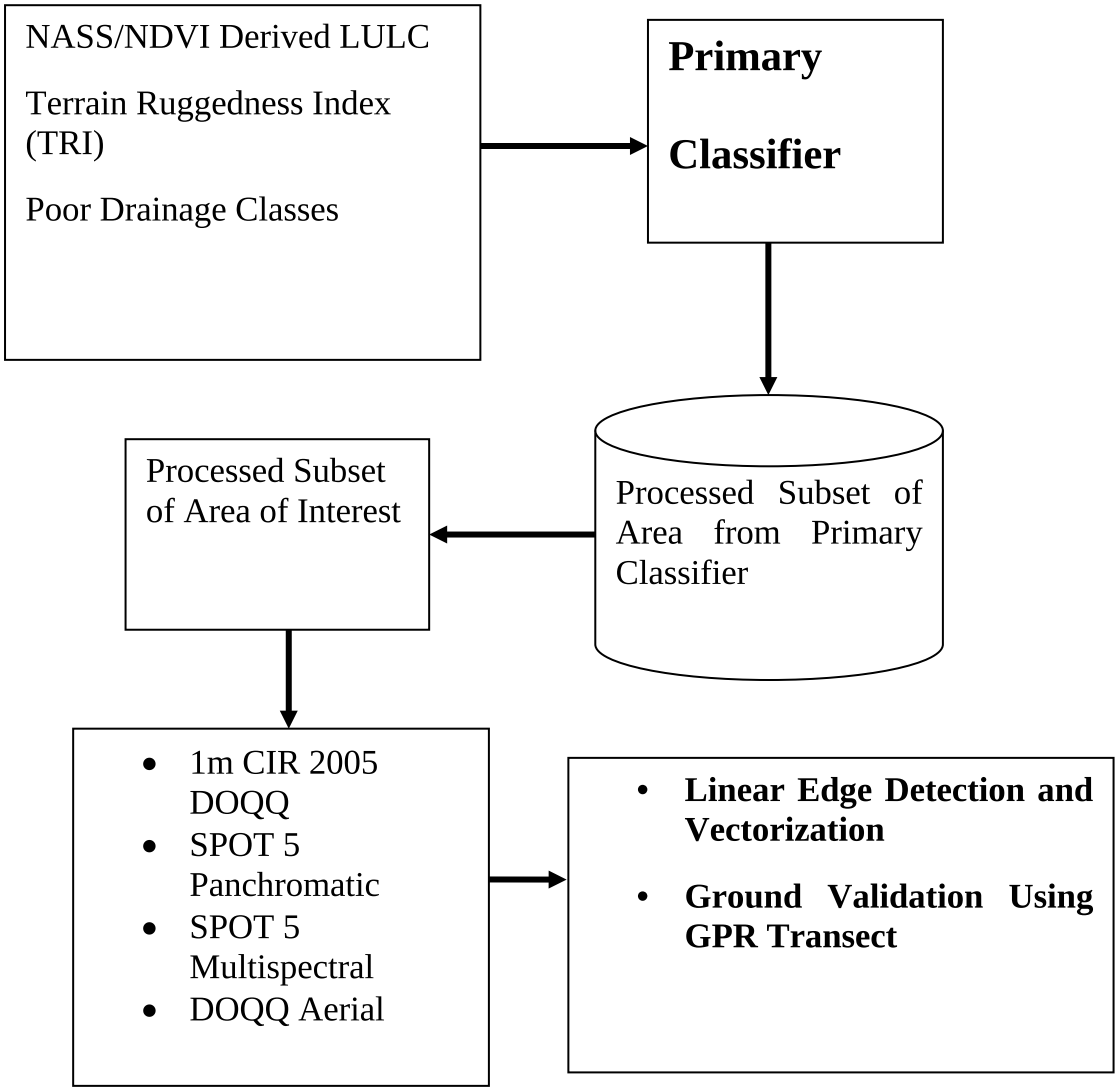

Fig. 5 Overview of data Analysis and processing

Fig. 5 provides an overview of analyzing the available datasets for detection of drainage tile, where primary classifier was first created for masking out urban and man-made structures from remote sensing images and different edge detection methods were then applied to masked images.

Primary Classifier

The primary classifier was derived from a combination of characteristics, which allow us to determine a more specific area of interest for tile drain detection. This stems from the consideration that the presence of urban or built up areas in agricultural watersheds poses limitations on the feature and terrain that can be delineated directly from remote sensing images, and that although most of the high resolution images were acquired during the leaf off period, the limitations from crop residue could be present. Land cover information, soil class data and topography are often used to form a decision classification tool (Varner, 2003; Naz, 2008). In this study, the creation of a primary classifier used three data layers: NASS/NDVI derived LULC, terrain ruggedness index (TRI), and poor drainage classes.

NASS/NDVI Derived LULC

The Land use and Land cover (LULC) data are obtained primarily from the National Agricultural Statistics Service (NASS) which generates this layer every year. The derived product is selected carefully for only agricultural classes so as to remove most of the impacts of built up areas. Since the data are derived from a coarse resolution dataset, the results are broadly validated by using a NDVI (Normalized Differential Vegetation Index) derived from high resolution CIR imagery.

Poor Drainage Classes

This dataset is based on specific soil characteristics which is dependent on the idea that poor drainage classes are the most prone to have tile systems installed. Hence the Soil Survey Geographic Database (SSURGO) soil data are utilized to extract dominant drainage class in the given area with poorly drained and very poorly drained areas which are used for our study.

Terrain Ruggedness Index (TRI) Layer

This layer was used to exclude high built up areas and areas of natural depressions such as streams and water bodies. It is one of the most critical indices and layers in understanding the attribute and the relationships between soil geomorphology and primary and secondary characteristics of flow in a given area. Primary attributes which are directly generated from elevation data include catchment area and point value measurements and also include derived datasets such as slope and aspect, whereas secondary attributes are linked with understanding more complex interrelations between the primary attributes and the spatial variability of specific processes occurring on the landscape. Terrain ruggedness index (TRI) is essentially topographic functions derived from primary soil attributes, and is established considering basic soil functions and characteristics shape of the terrain over a period of time (Moore et al., 1993). However, derivation of such layers was completed through a complex chain of processes as discussed below.

For understanding the variation of soil pattern with landform, a catenary landscape model called the compound topographic index (CTI) or the steady state wetness index, a secondary topographic attribute (Eq. 1) is often calculated from primary attributes such as slope (%) and specific catchment area to explore correlations between these attributes and the landscape hydrology processes (Gessler et al., 1995).

$$CTI = ln\left(\frac{A_S}{\tan\beta}\right) \qquad \text{(Eq. 1)}$$

where A_s is the specific catchment area(m^2 per unit width orthogonal to the flow direction) and β is the slope angle calculated in radian from the Digital Elevation Model (DEM). Owing to the limitations of CTI based approach for a flatter terrain and the strong correlation between CTI and TRI, the TRI layer serves as a more acceptable index for generating the primary classifier.

Terrain ruggedness index (TRI) is define as the largest inter-cell difference of a central pixel elevation to its surrounding cell and allows the segregation of hydrological parameters based on terrain relief (Riley et al., 1999). Here terrain ruggedness index (TRI) is computed as the difference between the value of a cell and the mean of an 8-cell neighborhood of surrounding cells. It can be summarized by the following expression

$$TRI = \sqrt{|(3x3max^2 - 3x3min^2)|} \qquad \text{(Eq. 2)}$$

where,

3x3 max = Focal statistical maximum value calculated using a 3x3 neighborhood function.

3x3 min = Focal statistical minimum value calculated using a 3x3 neighborhood function.

CTI and TRI prove to be important predictors and evaluators as they combine the contextual and site information along with primary attributes into detecting flow characterization and drainage characteristics over the given landscape. TRI can be used by extension as a hydrological predictor and marker, and is well validated and supported for implementation in the current study. The low ranges of TRI represent areas which are relatively flatter or have been modified to an extent such that they are low lying in some sense. Hence it is possible to use the TRI to suitably eliminate man made or natural depressions of large magnitude such as water bodies. This makes the dataset more applicable to a desired hydrological function and/or soil attribute derived data product, and allows the primary classifier to be less segregated and the resultants values are then aggregated by polygon aggregation method to generate a more efficient primary classifier. TRI thus serves as a more suitable index for analyzing areas of interest with low terrain variation (Mukherjee et al., 2012). The current model (Fig. 5) allows for the easy delineation of areas of interest in terms of topographical and morphological variations. The NASS/NDVI derived LULC is primarily masked over poor drainage class, and the intermediate raster image is then overlaid over the extracted TRI layer. This led to the creation of the raw primary classifier (Fig. 6 and Fig. 7) which is used as the mask for base imagery.

Fig. 6 Compound Topographic Index-Based Primary classifier

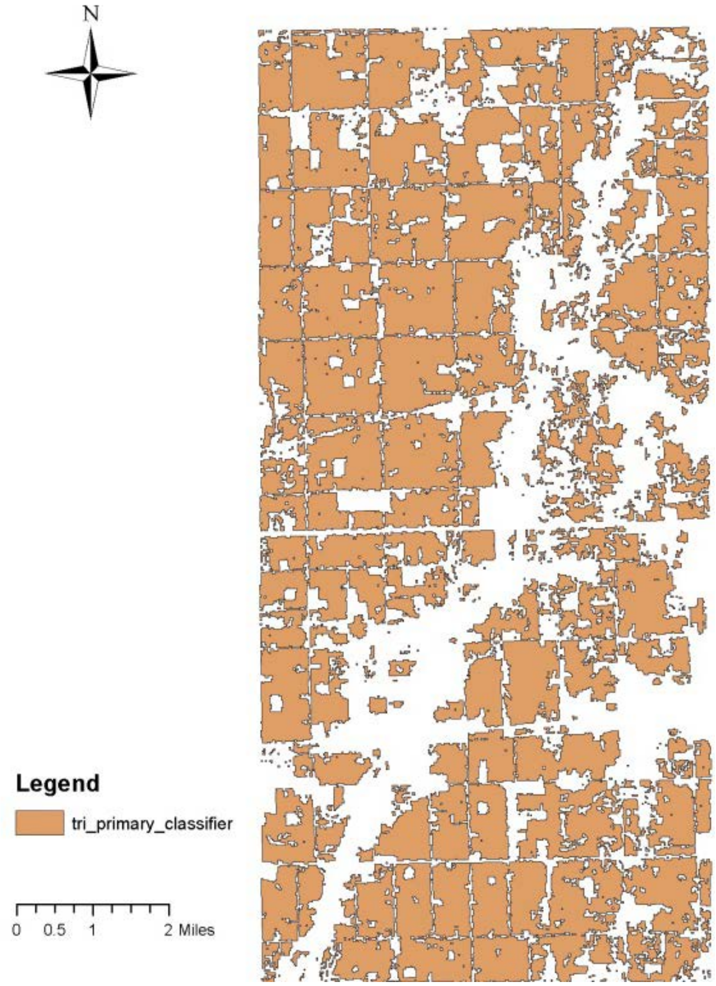

Fig. 7 Terrain Ruggedness Index-Based Primary classifier

Since the masked based imagery by the raw primary classifier was processed as a mask layer, it was then converted to a feature polygon and aggregated to include holes and minimum linear distances. Therefore the final mask was developed from the aggregated feature. The CTI derived primary classifier layer is retained for visual comparison with the TRI derived classifier mask.

The results clearly benefits from such a classifier since it eliminates built up, depressions or large water bodies such as the central stream in the given image segment. The developed classifier is novel, robust and easily applicable to an area with multiple land use classes which has a tendency to skew the results once edge detection methodologies are applied. It is important to understand that this step was quintessential in the process since it allowed us to reduce computational load and to improve primary edge detection from these sections.

Edge Detection and Image Segmentation

The application of edge detection in remote sensing has found a major niche for conventional agricultural inventory and for feature detection in major agricultural watersheds. Edge detection has been one of the most commonly used methods for such studies along with the identification of features, which is now possible owing to high resolution imagery. The methodology for edge detection is often dependent on the idea that variations in pixel intensity, saturation or color lead to a break in a signal which can be identified in an edge. It has been discussed by researchers in the field that remotely sensed imagery is a broadly thus a discrete representation of spatio-temporal magnitude of reflected energy (Rydberg and Borgefors, 1999). The classic case of edge detection using the principle of signal theory identifies an edge as the point where there is rapid shift of signal or pixel value between the neighboring cells. A variety of edge detectors are developed. Since the sources and the type of imagery are varied, the methodology and the efficiency of a detection method vary.

Kernel Based Edge Detection Techniques

The basic edge detection algorithms (Roberts, Sobel and Prewitt) consist of matrix or kernel based gradient operators. For example, the Roberts filter is a 2x2 matrix operator, while the Sobel and Prewitt operators are standard 3x3 gradient edge detectors. In the operation of these filters, convolutions are performed primarily on the center cell for the convolved matrix of image data (Sun, 2012). The new value of the central pixel replaces the old values, and the filtered image after convolution is then run through a given threshold to decide if it is an edge. These gradient based algorithms use kernel operators to calculate edge strength along horizontal and vertical direction. The contribution of these different components is then combined to give the total value of the edge strength. For the Sobel filter, the two kernel matrices are

$$Vertical\ matrix \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

and

$$Horizontal\ matrix \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}$$

The Sobel operator is operated in both directions, and the derived components are added to generate the final strength of edges in the imagery. The Sobel filter was applied to CIR, natural RGB and DOQQ imagery.

Similar to Sobel, the Prewitt operator measures the vertical and horizontal edges with kernel function and then gives the intensity differences of the gradient in the current pixel. Again in this case the horizontal and vertical matrices are

$$Vertical\ matrix \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}$$

and

$$Horizontal\ matrix \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

Both these methodologies can detect long continuous edges and are sensitive to noise. Both Sobel and Prewitt filters were applied to the base image to discern the delineated edges and the pattern of detection.

The kernel size and the signal to noise ratio in an image play important roles in detection efficiency. Smaller kernel size and the noise in the imagery limit the effectiveness of these detection techniques. An improved User defined directional kernel in rectangular is introduced. This User defined kernel is well suited for area with low variations in topography and consists of a 5x7 horizontal convolution matrix and a 7x5 vertical convolution matrix.

$$\text{Horizontal Matrix} \begin{pmatrix} -4 & -3 & -2 & -3 & -4 \\ -3 & -2 & -1 & -2 & -3 \\ -1 & -0.5 & -0.25 & -0.5 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0.5 & 0.25 & 0.5 & 1 \\ 3 & 2 & 1 & 2 & 3 \\ 4 & 3 & 2 & 3 & 4 \end{pmatrix}$$

$$\text{Vertical Matrix} \quad \begin{pmatrix} -4 & -3 & -2 & 0 & 2 & 3 & 4 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -2 & -1 & -0.5 & 0 & 0.5 & 1 & 2 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ -4 & -3 & -2 & 0 & 2 & 3 & 4 \end{pmatrix}$$

The horizontal matrix has smaller kernel element step size (0.25), because the finest edges are visually detectable along horizontal directions of the image. For this study, the process is automated using the Model Maker functionality of ERDAS IMAGINE whereby the process can be run for all the masked imagery. Since the visibility of detected edges requires the scale of the image to be smaller means, a scaled or zoomed in subset of the main imagery can be utilized to show the actual edges that have been delineated. The filtered image is then run through an ISODATA unsupervised classifier. In this process where an arbitrary initial class is assigned followed by classification of each pixel to the closest pixel until a user defined threshold value is reached. In the current case, the threshold generally refers to percent refinement by splitting and merging of clusters and to retain the number of classes for the unsupervised classification. The method creates cluster for edges which can then be converted into vector data using ArcGIS based vectorization. These vector layers are used for intersect analysis.

Combination of Receptive Fields (CORF)

The idea of understanding image segmentation as a function of the visual field called the receptive field has been explored for quite some time. It is based on the fact that a neuron would respond to an edge or a line of a given orientation in a given area which is the receptive field (RF). The human cortex visualizes multiple fields in an image while differentiating edges and segmenting object. The human visual cortex benefits from 3 dimensional field of view compared to the 2D image segmentation in natural vision, and are also dependent; not only on contrast response but also the effect of optimally oriented stimulus. The CORF model allows the user to incorporate and exhibit cross orientation suppression, contrast invariant orientation tuning and also response saturation (Azzopardi and Petkov, 2012).

Since functioning as a contrast and edge delineator for field cells or neurons, CORF is an excellent method for delineating natural contours that exist in imagery. The procedure includes edge thinning by non-maxima suppression followed by binarization using hysteresis thresholding. However the growth of the detection sensitivity to the stimulus contrast is not proportional, meaning there is the response saturation when it comes to edge detection in imagery. The CORF orientation model allows for the response saturation which is more realistic to real simple cells. The CORF model has a good response to natural contour characteristics.

A MATLAB code has been adapted for this purpose and was run on all of the images for a comparison with the other edge detection methodology and to understand the limitations of such algorithm. The images are resized owing to the limitations in the memory that is used by the program. A section of the imagery has been enlarged and resized to understand the effect of using imagery at different scale and the capability of the algorithm in resolving edges at varied scales.

As already discussed the CORF based detection is also dependent on characterization of texture and natural contours of neurons apart from simply differences in intensity and thus is beneficial in some sense to the proposed problem. CORF has limitations to discern finer details and only delineates the external boundary of major natural contours. This method works better on smaller scale edges owing to the thinning algorithm. The limitation of CORF can also be caused by the fact that the imagery has crop cover and other regular pattern which induce extra interference for edge detection.

Phase Congruency Detection (PCD)

Phase congruency detection (PCD) is a corner and edge detection operator which uses the principal moments of the phase congruency information to determine the corner and edge. The edge detection technique inculcates the fact that the results of edge detection are highly localized and thus are invariant to image contrast (Kovesi, 2003). The method isolates both edges and corners, and benefits from the fact that the corner map is a subset of the edge map.

A corner for PCD is defined as the location in the image where there are distinct peaks for the local autocorrelation. Most corner detection algorithms are sensitive to image contrast which increases if the sequence of images or the area covered by the images is large. The change in image contrast limits the setting of a threshold for the same image. The corner detection is also impacted by the fact that most smoothening algorithms such as Gaussian smoothing corrupt the location of corner (Kovesi, 2003).

It has been found early on that for any wave form, the highest peak or corner is where the Fourier components lie exactly in phase in the phase congruency model, and that while the shape of the waveform is not always known, we can simply look for the point with high degree or arrangement/order in the Fourier domain. Locating phase congruency is complex but this can be done by locating the peaks in the local energy function (Eq. 3). The local energy function is defined for a one dimensional luminance profile, F(x) as

$$E(x) = \sqrt{F^2(x) + H^2(x)} \qquad \text{(Eq. 3)}$$

where, H(x) is the Hilbert transform of F(x) (a 90 degree phase shift of F(x))

It can be shown that energy is equivalent to phase congruency which has been scaled by the sum of the Fourier amplitudes (Eq. 4)

$$E(x) = PC(x)\sum_n A_n \qquad \text{(Eq. 4)}$$

Since the local energy function is proportional to the phase congruency function, the peaks in local energy could thus be related to phase congruency. As a result the local energy could be calculated by convolving the signal with a pair of filters in quadrature (Venkatesh and Owens, 1989; Kovesi, 2003). For the purpose of our work, the complex valued Gabor function is used with has a sine and a cosine wave modulated by a Gaussian function.

A MATLAB code was adapted for this purpose and has been run on all of the images for a comparison and to realize the importance of image scale and limitations of using such a process. Owing to the limitations in processing memory required by MATLAB, the images were resized and rescaled. PCD performs differently for images at different scales owing to change in the energy function at different scales. Sections of the imagery were selected similar to the CORF model for running the detection and for resolving edges and corners.

Ground Validation and GPR data analysis

The GPR dataset provides important information on the location of tile outlets in addition to the Baker tile clip and was available to us in terms of point transect data for the GPR transects and as polyline features for the Baker tile clip. The GPR transects are used to validate and assess the usability of the defined processes for the tile drain detection. The area of interest with the GPR point transects is thus scaled, and major comparisons are made both visually and with the help of intersect analysis between the point & polygon outlet layer along with the delineated drainage vector layer.

RESULTS AND DISCUSSION

The current study results are categorized based on the type of imagery and the methodology. The variation of data types occurs along two distinct domains, the spatial and temporal resolutions of the images being used. The spatial resolution is important because the features can only be delineated if they are within a given range, and the temporal resolution is relevant to the time period after rainfall during which the image is collected and to the interference with the tile detection of the crop residue that might be left on the field.

SPOT 5 (Pan Chromatic and Multispectral)

The SPOT 5 dataset consists of the pan chromatic data with the highest resolution in the dataset of about 5m and the multispectral data with a spatial resolution between 10-20m. Both panchromatic and multispectral data were analyzed using Robert, Sobel and Perwitt and a user defined kernel. The SPOT 5 imagery did not produce viable results for Roberts or Perwitt methodology, but produced discernible results for Sobel filter. The image symbology was changed and the darker shades of green and brown indicated the detected edges and are blurred similar to optical aberrations for both panchromatic and multispectral datasets (Fig. 8 and Fig. 9).

Similarly, the user defined filter could not improve the quality of the edge detection, which is expected since the artifacts and the noise to signal ratio are high. Therefore both panchromatic and multispectral SPOT 5 datasets were not suited owing to the fact that major feature sets were not discernible.

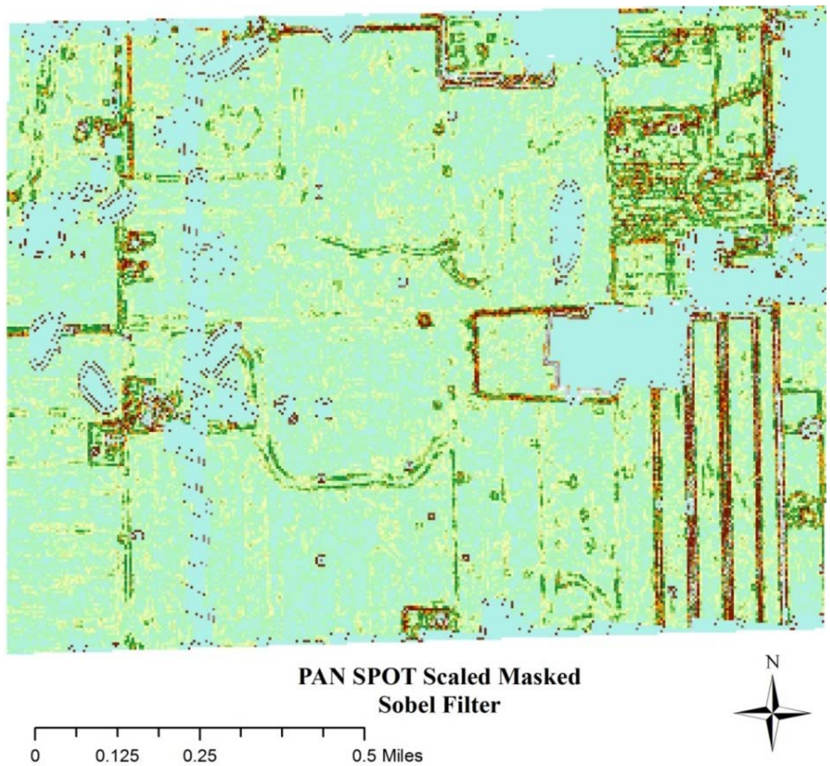

Fig. 8 Panchromatic SPOT 5 Processed through Sobel Filter

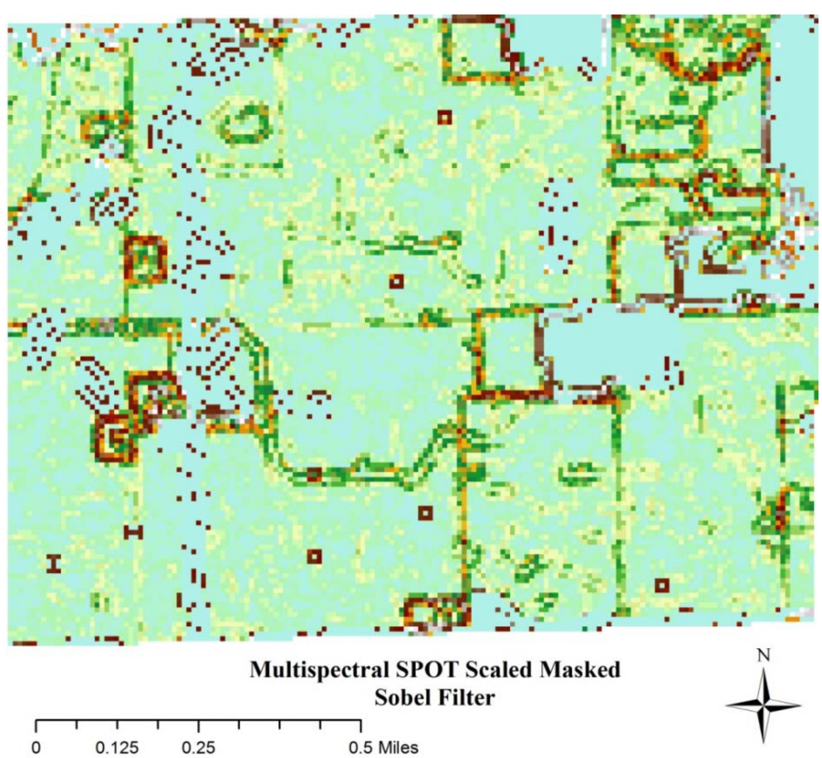

Fig. 9 Multispectral Spot 5 imagery processed through Sobel filter

CORF and PCD which rescale the data, average the pixel intensity and contrast further, degrade the actual pixel value and produce less than average results when applied to SPOT datasets, which were modified by the primary classifier. Most of the kernel-based methods did not perform very well for the SPOT images owing to a large noise to signal ratio and the smaller kernel sizes. Since a lot of these images have existing noise and crop cover residue that existed for the region, understanding and implementing an appropriate kernel for filtering was complex.

Color Infrared Imagery

Color infrared imagery has 1m spatial resolution and was expected to perform better than the SPOT 5 datasets. The best results are obtained from the Sobel filter (Fig. 10) for the standard kernel edge detection and by using the user defined kernel (Fig. 11).

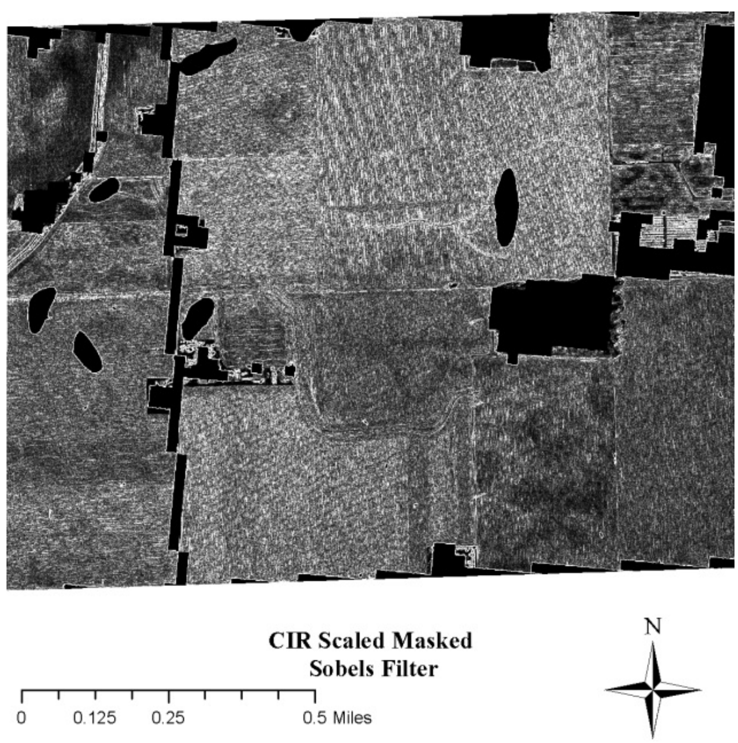

Fig. 10 CIR Scaled subset area using Sobel filter

As shown in Fig. 10, the Sobel filter is effective in detecting large variations in intensity and is capable of resolving the edges. However, the detected edges along the drainage patterns seem to be patchy and segmented. This is due to the crop cover residue as well as soil moisture content which tend to skew the detection. The user defined kernel filter was able to get rid of minute artifacts present in the imagery and produced more refined results for the same scaled subset data.

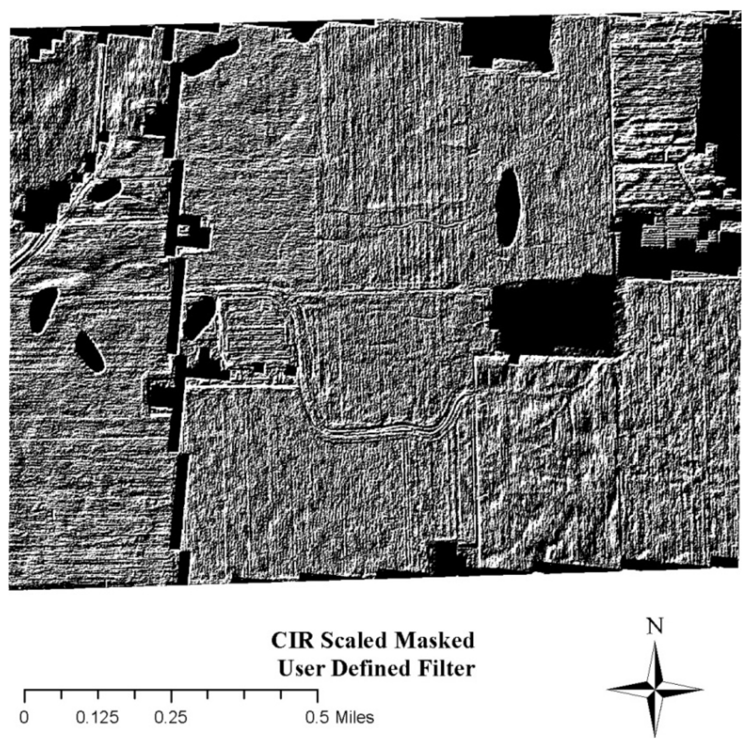

Fig. 11 CIR Scaled subset area using user defined kernel filter

Digital Orthophoto Quarter Quads

The DOQQ imagery represents one of the best combinations of high spatial and temporal resolution. The imagery was collected from airborne sensors and had an average resolution of about 1m which is aptly suitable for the detection of our current tile features. The dataset is also known to have been collected after a major rainfall event and had hence benefitted the current analysis. For this data, the best standard kernel methodology was available from the Sobel filter whereas better results were obtained from the user defined kernel (Fig. 12). Additionally, distinctly better results were obtained by PCD and CORF. The results obtained clearly elucidate the importance in the choice of image datasets in conjunction with the appropriate methodology.

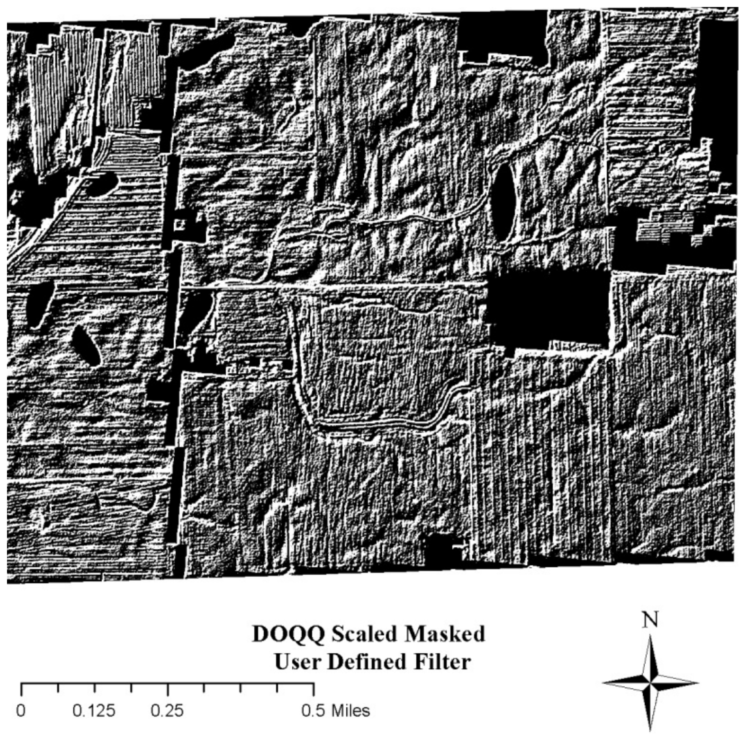

Fig. 12 DOQQ Scaled subset area using user defined kernel filter

Since the best results were obtained for the DOQQ imagery, it is necessary to provide an insight into the output for this image input. To achieve this, the results for this dataset were compared for the efficiency of the different methodologies. The comparison is based on the ground GPR validation data provided by USGS. Two clips are compared for the given area of interest, the Baker tile clip which is clipped to a polygon consists of ground validated tile drain features which are polyline features. This can be referred to as polyline validation clip and the next clip which consists of point GPR tile outlets are referred to as point validation clip. Both polyline and point validation clips were chosen to keep into consideration the availability of the tile features for comparisons along with the ground validation data.

For the proposed method, a primary classifier mask was used to create an aggregated subset area which was scaled and used as the standard area for comparison of the performances of the standard detection methodologies and the overlay with the Baker tile clip for the same area was assessed. The standard kernel methodologies as discussed earlier included Roberts, Perwitt and Sobel.

The Roberts filter seems to produce minutely discernible filtered edges and is not beneficial for detection of tile features. It used a 2x2 matrix which was not optimal since it relied on the continuous gradient at the interpolated point. It was not as an effective edge detection filter since natural imagery consists of irregular edgings and has a tendency to contain a large amount of noise coupled with image artifacts. However, both Sobel and Perwitt consist of a 3x3 matrix directional filter, which are effectively better than the results from smaller matrix sizes.

Because of the constant step or operator element for Perwitt, it does not provide an emphasis on the central pixel in the same way as the Sobel filter and the former did not generated as good results as those from the latter. Sobel allowed for the removal of minute artifacts in the images, thus improving edge detection. The interpreted tile was digitized and overlaid (Fig. 13) while an intersect analysis provided the overlay length percentage (Table 8). This reiterates the fact that edge detection by Sobel was not automated, and manual digitization was performed to delineate edges from the filtered image. It is expected that the user has a tendency to introduce his/her interpretation and bias for a specific feature. The scaled clip was also utilized for other developed methodologies and this will be discussed sequentially in the succeeding paragraphs.

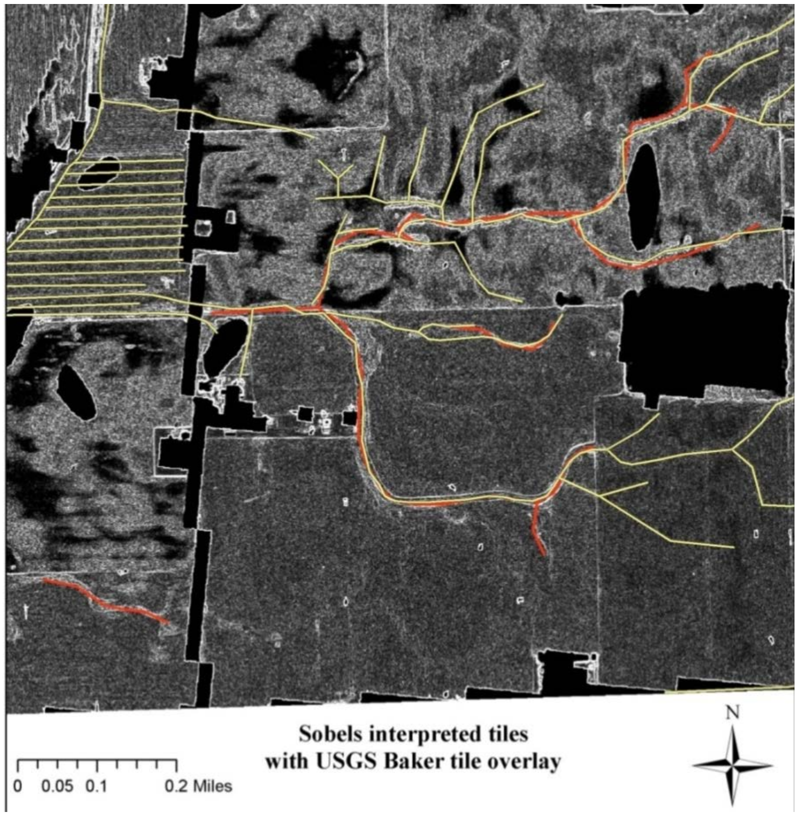

Fig. 13 Sobel interpreted tiles with USGS Baker tile overlay

An overlay accuracy of only 18.28% resulting from Sobel has been used as a representative accuracy (Table 8) for the performance of 3x3 operators for edge detection, and Sobel had a low commission error (4.22%) but an extremely high omission error of about 81.77%. This clearly represents the fact that the Sobel derived results were not suitable in capturing majority of existing tile network and Sobel is hence not an optimal method for tile drain detection.

Table 8. Tile Overlay (% overlay of tile length) & error estimation

Length	Total Length(m)	Intersect Length(m)	Overlay (%)	Commission Error	Omission Error
User Defined Intersect	13322	13322	81.68	0	18.31
PCD Intersect	43961	11487	70.43	199.11	29.56
CORF Intersect	19490	6778	41.55	77.94	58.44
Sobel Interpreted	3671	2982	18.28	4.22	81.71
Baker Tile Clip	16309				

The results for the standard kernel filters depicted and strengthened the fact that spatial heterogeneity of the given landscape and natural profiles create a complex environment for analysis (Sun, 2013). The limitations in the kernel size of the standard filters and the complexity of the natural landscape and features emphasize the development for a user defined kernel. A user defined kernel performs with higher efficiency for the same given area. The overall overlay percentage of over 81% indicates (Table 8) that the user defined kernel serves as the most efficient edge detection technique and is independent of user bias during digitization. The user defined kernel benefits from its rectangular kernel, which allows edge detection to be more prominently extended in a particular filtering direction. The user defined kernel uses smaller step difference within the kernel elements, which allows for capturing minute variations in edges efficiently along both horizontal and vertical direction and provides better functionality in flatter topography.

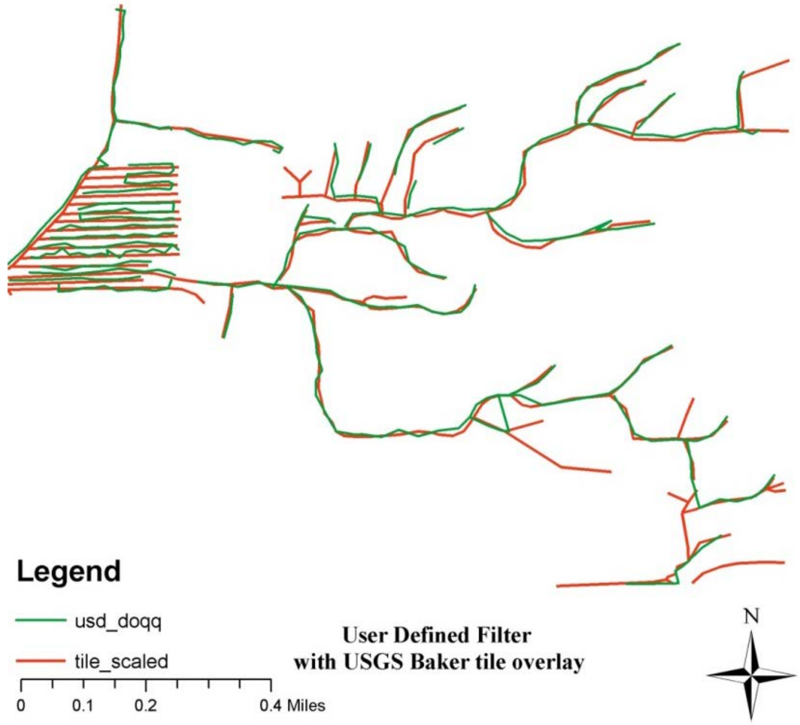

Fig. 14 User defined filter with USGS Baker tile overlay

As expected the user defined kernel had negligible commission error and lowest omission error of about 18.31% (Table 8). This clearly justifies using this methodology for edge detection and tile layout delineation, and establishes itself as the most efficient among all applied methodologies described so far.

While the limitations in the standard kernel based filters (both gradient and non-gradient based filters) emphasized the development of a better kernel methodology, the limitations also motivated the development of certain natural vision based edge detection methodologies that were suitable for interpretation of edges in natural imagery. As mentioned in section Edge Detection and Image Segmentation, CORF is an image detection and filtering methodology based on natural feature detection mimics properties with real simple cells in vision. The model allows for detection of features such as cross orientation suppression and response saturation and works better in terms of detection of saturation invariant responses.

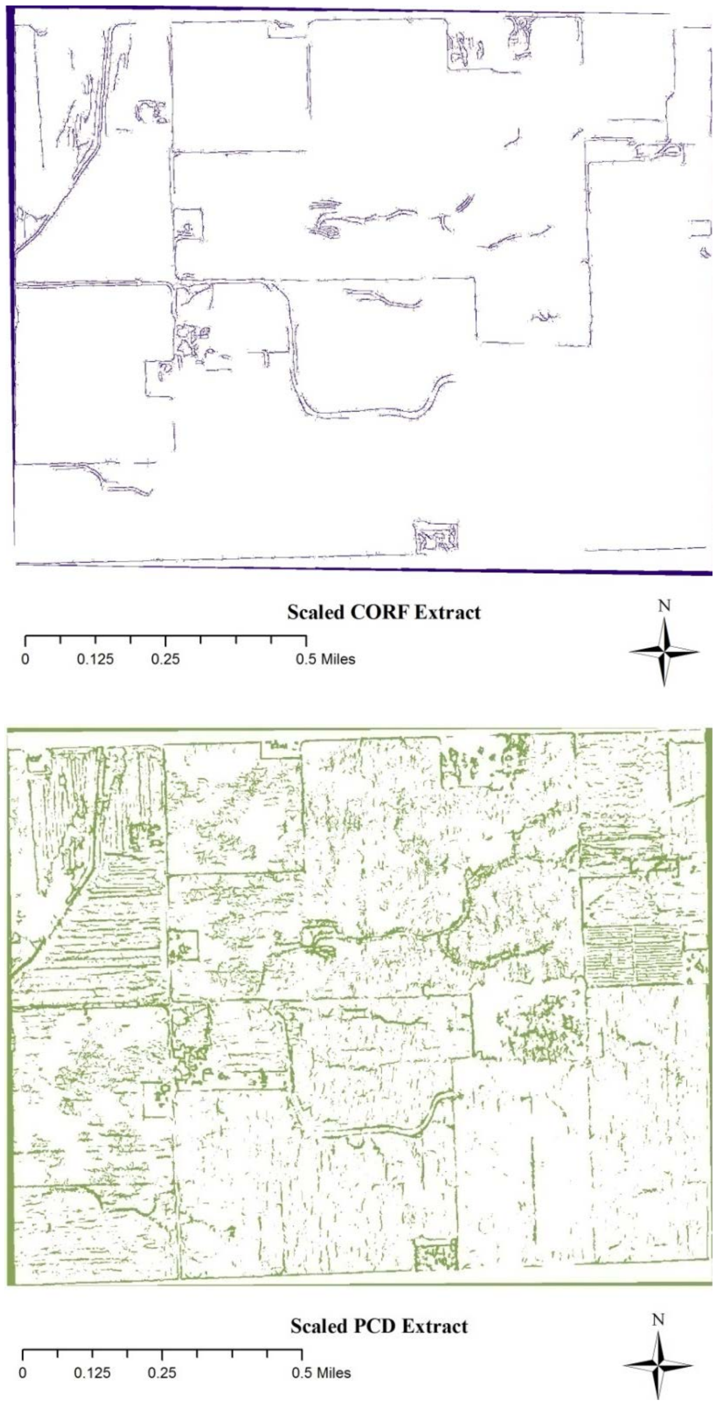

Fig. 15 CORF and PCD edge extract for Scaled clip

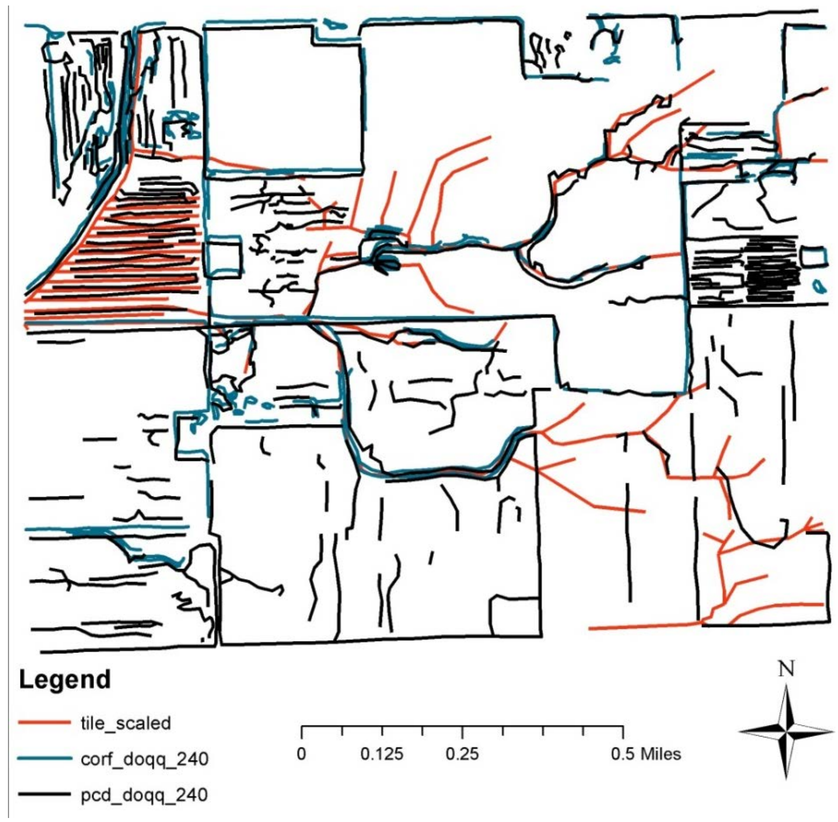

Fig. 16 CORF and PCD extracted tiles with existing tile layer overlay

As shown in Fig. 15, results for both CORF and PCD were dependent on intensity and contrast variations in imagery, coupled with both edge as well as corner detection. The overlay results for PCD and CORF are shown in Fig. 16. The results demonstrate that although CORF, a contrast dependent methodology, was not well suited for detection of weak edges in the imagery, the result has an accuracy of over 42% better than the existing standard kernel methodologies (Table 8). The computational limitation to CORF for handling large images requires rescaling, and the rescaling step reduces the number of sharper contrast variations and the intensity that creates the stronger and weaker edges. The CORF results also had a commission error of 77.94% meaning a large segment of the detected result did not overlay with the Baker clip; the omission error of about 58. 44% (Table 8) for this method also emphasizes the failure to capture a large part of the tile layout network. While thinner and stronger edges are detected more accurately, it does not work well with low intensity weaker edges or with low contrast variations in the edges. Nonetheless, the current work is one of the preliminary applications of natural vision based edge detection in remote sensing and serves as an outline for further work to be continued which could include intensity variance along with contrast variances in both small scale and large scale imagery.

Phase congruency detection (PCD) is another adaptive methodology which generated a higher efficiency of over 71% and captured both weak and strong edges. The result for PCD also shows that its weak edge detection capacity (Fig. 15) leads to over detection of edges. Therefore PCD had a large commission error of over 199.11. An omission error of 29.56% however ascertains that the method does capture in some sense most of the tile segments from the Baker clip indicating that it is more efficient (Table 8) and better than CORF.

However, PCD's dependency on the analysis window contributes to the local frequency components to be present and analyzed. This limits the functionality of the PCD detection which is based on the scale of the imagery and the analysis window in conjunction to their surrounding features/environment (Kovesi, 2003). Kovesi (2003) suggested that a smaller analysis window could contain multiple features with a great degree of independence from other features and hence each feature was perceived to be more important locally with a better response. In our case the feature sets being detected are fine resolution datasets and hence their detection over a smaller spatial scale allows the PCD to perform with higher output. Thus the performance of PCD is image scale dependent and the result could be better when the feature of interest and the scale of the imagery are closely related. Thus both CORF and PCD have functional limitations which can be improved further as the methodology is better adapted for the given type of imagery.

GPR Validation

The Baker clip overlay, over the polyline validation clip with an overlay analysis ascertained the efficiency of a developed kernel and natural vision based methods. The GPR outlet transect points were used to validate the placement of tiles (Fig. 17) which drain into the outlets and can be interpreted by the overlay. The GPR point outlets in the point validation clip were overlaid for section of the imagery and the PCD, CORF and user defined kernel results were evaluated for their overlay (Fig. 18).

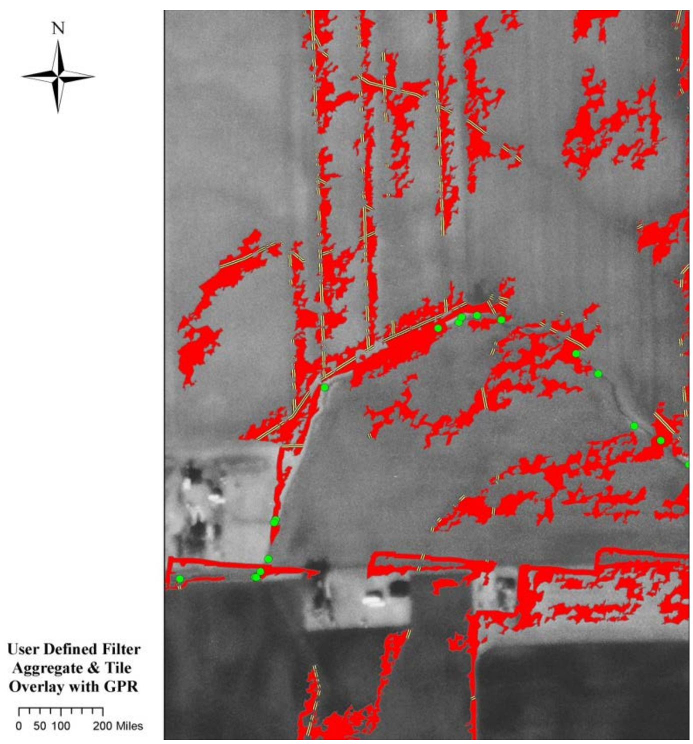

Fig. 17 User defined filter aggregated with GPR outlet overlay

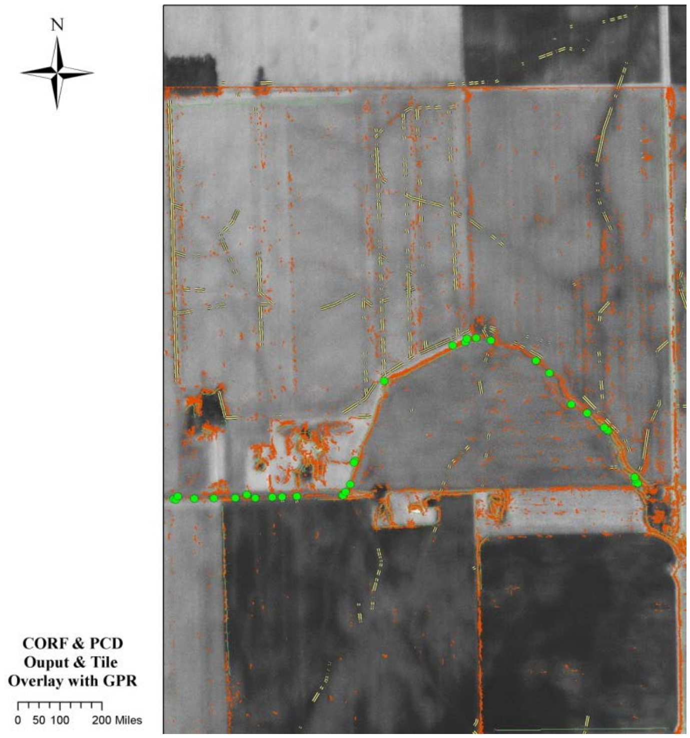

Fig. 18 CORF and PCD tile overlay with GPR outlet overlay

Optimal combinations of methods with datasets

Our current work explores the extraction of tile drain systems keeping in mind the limitations of both the datasets used and the methodology. Understanding optimal combination of the datasets along with the methodology used provides an insight and a guideline into tile drain detection. The entire study focuses on determination of the characteristic combination of the data type used and the methodology that has been applied which is summarized in the table below (Table 9).

It is clearly depicted from the results that spatial resolution of the imagery plays an important role in feature detection. This is because of the fact that finer features such as tile drains are only discernible in high resolution imagery. The SPOT imagery had varying spatial resolution from 5m for panchromatic and 10-20m for multispectral image.

As expected at this resolution application of kernel methods and even user defined kernel and vision based methodology do not function well.

Table 9. Image and methodology selection for edge detection analysis

	Spatial Resolution	Kernel Based Methodologies	User Defined Filter	CORF & PCD
SPOT 5 Pan & Multispectral	5m 10-20m	Sub Optimal results, low spatial resolution	Sub Optimal results, edges blurred	Rescaled, results sub optimal
CIR	1m	Sub Optimal results with high crop cover residue	Sub optimal results with over detection of edges & artifacts	Rescaled, results suboptimal with over detection of edges
DOQQ	1m	High spatio-temporal resolution, over detection of noise	Best Optimal solution, suited for removal of noise	Fairly optimal results with variability with weak & strong edge detection results

The choice of spatial resolution was hence stressed in the next pair of imagery which were CIR imagery from NAIP and the CIR imagery from USGS aerial photography. Both of the images had high spatial resolution of about 1m and were expected to generate better results. The CIR imagery from NAIP even with high spatial resolution had large crop residue owing to the time period or temporal resolution of the imagery. The crop residue makes the reflectance from these tile drain non discernible and hence the NAIP CIR imagery even with its high spatial resolution could not function very well with the available filtering methods. The crop residue also generated artifacts and makes the image noisy and unsuitable for filtering operations. As expected the Sobel filter performed the best among known filtering kernels while the user defined kernel, CORF and PCD had a tendency to over detect false edges owing to the image characteristics.

The choice of imagery has thus been a optimal balance between the spatial and temporal resolution by itself. The best choice of imagery for such a study would include high spatial resolution 1m or less and should be taken post a rainfall or a storm event allowing tile drains to be activated and the drainage patterns to clearly appear in the imagery. As expected the temporal resolution of the imagery plays a crucial role in the visibility of activated tile drains and the data might hence vary with varying intensity of rainfall and the drainage period. In our case the DOQQ as expected performed crucially well fitting into both the spatial and temporal domains that were necessary and was hence chosen as the base imagery.

The next step in understanding optimality condition includes understanding the combination of the image type with the edge detecting algorithm. The DOQQ imagery obtained from USGS was used for the benefits in terms of lesser noise and high spatial and temporal resolution. The DOQQ imagery works sub optimal with the existing kernel methods such as Roberts, Sobel and Prewitt. As discussed earlier the smaller kernel size limits the performance of edge detection and hence course resolution imagery (SPOT 5) as well as high spatial resolution imagery such as NAIP CIR and DOQQ do not perform well. The combination of DOQQ performs best with user defined kernel, which is the optimal combination since it allowed for detection of finer edges and reduced noise. CORF and PCD on the other hand work fairly optimal but have reduced efficiency in terms of commission and omission errors that they encounter.

Having an optimal solution however raises questions about the existing limitations in methodology, making a given solution a robust best solution apart from being locally optimal.

CONCLUSIONS

The subsurface drainage structure existing in the Little Sugar Creek-Sugar Creek watershed has been the target for testing the efficiency of different tile detection and layout methodologies. Differing from previous studies, the work moves from understanding the layout of regular and irregular tile drain systems and focuses on the limitations of earlier methodologies. A primary classifier system was developed which takes into consideration topography in terms of TRI, the vegetated area from the LULC and the soil classification layer. The classifier allows us to extract only the area of interest and to effectively mask out the remaining area. The step allows an overall improvement in computational resource for the study while preventing the results to get skewed. Several filters and operators were compared including Roberts, Perwitt and Sobel along with a user defined filter and CORF and PCD methodology.

The limitations in kernel size causing high noise to signal ratio was resolved with the development of a novel user defined kernel. The developed kernel included a rectangular kernel size which was more suitable for directional edges and smaller element size step was chosen to be able to filter out weaker as well as stronger edges. As a result the user defined kernel produced the best optimal results with extremely low omission error and seems to be promising being an automated methodology. The motivation for natural vision based edge detection was drawn from observing how corner and edge detection by natural vision is more efficient and is independent of orientation. Both CORF and PCD produced fairly optimal results with issues of weaker and stronger edge detection attributed to image rescaling. This was the first time that these methods had been adopted to the domain of remote sensing dataset analysis and substantial improvements can be expected in analysis methodology.

Optimal dataset and methodologies were compared and tabulated for the efficient use of applied edge detection algorithm. The user defined filter benefited the results largely owing to its large kernel size and smaller kernel element step. The results also pointed out that the low omission and commission error were preferable for this situation. The CORF methodology on the other hand was found to be effective only for stronger edges and were limited by the rescaled image. Since the rescaling cause loss of image intensity and contrast differentiation there is loss of actual edges that could be discerned. Thus CORF produced fairly optimal results for tile drains with stronger edges. PCD was a more efficient natural vision based methodology but owing to sensitivity to both strong and weak intensity and contrast, over detection was seen in the imagery.

The novel user defined kernel along with CORF and PCD suffer from limitations of existing operator systems that are used for edge detection. The input resolution and the time period of the imagery were identified as important factors which were evaluated for determining actual edge detection efficiency.

The comparison of user defined kernel, CORF and the PCD methodology elucidate and focus on the following points and can be summarized.

- Both CORF and PCD owing to the limitation to image size can benefit from smaller subsets which are processed sequentially and then mosaicked for the area of interest.

- The idea of automation to the tile detection is important and can be achieved easily from all of the methods there were developed and adapted for the current study.

The developed methods along with CORF and PCD are more flexible in terms of area of application and should benefit the tile layout and detection efficiency further. The study thus deals with the problem of suggesting techniques for tile layout detection which are easily adaptable and highly efficient for our given situation.

FUTURE WORK

The best tile detection scenario inculcated optimal combination of both datasets and the used edge detection methodology. Limitations in accessibility of the aforementioned data and the related time frame in which the data is obtained results in a fairly narrow window. The frequency of data collected after a rainfall event of varying intensity could allow us to cross validate the same tile systems. This will also create a continuous data segment which can be recorded during the rainfall seasons to map functioning of tile drains. For an efficient edge detection this must be in combination with the methodology that is being used, in our case the user defined filter performed best and needs to be tested for establishing a robust methodology over the given area of interest.

There is a need for periodical data for an analysis of this nature; which will allow the analysis to pose solutions to existing limitations and to be able to generate a more efficient detection tool. Secondary aspects of tile drain detection also include the aspect of tile intensity (no of tiles/area) and the tile incision depth which correlate to the productivity in some sense. A predictor system can thus be developed further as the quality and time period of collected data increases, allowing prediction models to comment on increase in productivity along with efficient tile drain placement.

Though there are clear limitations in terms of spatial and temporal resolution of the datasets, the developed methods and workflow allow for automatic delineation of the tile drain patterns which was our primary objective. The CORF and PCD methodologies need to be modified with more computational capability, for larger dataset while preserving the geo-referencing information. The future work must also resolve the problem with over detected and under detected edges and must allow for secondary analysis such as continuity analysis or junction analysis for such a connected network. . Integrating spacing between tiles and junction analysis could be utilized for eliminating false edges. An effective study keeping in mind the multiple criterions that enter into the detection and delineation of layout is complex and can be improved with improvement in datasets and the methodologies developed further.

REFERENCES

1. Allred, B. J., Fausey, N. R., Peters, L., Chen, C., Daniels, J. J., & Youn, H. (2004). Detection of buried agricultural drainage pipe with geophysical methods. Applied Engineering in Agriculture, 20(3), 307-318.
2. Azzopardi, G., & Petkov, N. (2012). A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model. Biological Cybernetics, 106(3), 177-189.
3. The Purdue Center for Climate Change Research (2008). Impacts of Climate Change for the State of Indiana
4. Cooke, R., Badiger, S., & García, A. (2001). Drainage equations for random and irregular tile drainage systems. Agricultural Water Management, 48(3), 207-224.
5. Franzmeier, D. P., E.J. Kladivko, B.J. Jenkinson. 2001. (2001). Wet soils of Indiana Purdue Extension Publ (Vol. AY-301).
6. Gessler, P., Moore, I., McKenzie, N., & Ryan, P. (1995). Soil-landscape modelling and spatial prediction of soil attributes. International Journal of Geographical Information Systems, 9(4), 421-432.
7. Goswami, D., Kalita, P. K., Cooke, R. A. C., & Mclsaac, G. F. (2009). Nitrate-N loadings through subsurface environment to agricultural drainage ditches in two flat Midwestern (USA) watersheds. Agricultural Water Management, 96(6), 1021-1030.
8. Kovesi, P. (2003). Phase Congruency Detects Corners and Edges. Paper presented at the Proc. VIIth Digital Image Computing Techniques and Applications, Sydney.
9. McKenzie, N., & Austin, M. (1993). A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation. Geoderma, 57(4), 329-355.
10. Moore, I., Gessler, P., Nielsen, G. A., & Peterson, G. (1993). Soil attribute prediction using terrain analysis. Soil Science Society of America Journal, 57(2), 443-452.
11. Moore, I. D., & Grayson, R. B. (1991). Terrain-based catchment partitioning and runoff prediction using vector elevation data. Water Resources Research, 27(6), 1177-1191.
12. Mukherjee, S., Mukherjee, S., Garg, R., Bhardwaj, A., & Raju, P. (2012). Evaluation of topographic index in relation to terrain roughness and DEM grid spacing 2. ISRO Internal.
13. Naz, B., & Bowling, L. (2008). Automated identification of tile lines from remotely sensed data. Trans. ASABE, 51(6), 1937-1950.
14. Oosterbaan, R. (1994). Agricultural drainage criteria. Wageningen, The Netherlands: Water Resources Pubns.
15. Pavelis, G. A. (1987). Farm drainage in the United States: History, status, and prospects: US Dept. of Agriculture, Economic Research Service.
16. Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. intermountain Journal of sciences, 5(1-4), 23-27.

17. Rozemeijer, J. C., van der Velde, Y., McLaren, R. G., van Geer, F. C., Broers, H. P., & Bierkens, M. F. P. (2010). Integrated modeling of groundwater-surface water interactions in a tile-drained agricultural field: The importance of directly measured flow route contributions. Water Resources Research, 46.
18. Rydberg, A., & Borgefors, G. (1999). Extracting multispectral edges in satellite images over agricultural fields. Paper presented at the Proceedings. International Conference on Image Analysis and Processing.
19. Schilling, K. E., & Helmers, M. (2008). Effects of subsurface drainage tiles on streamflow in Iowa agricultural watersheds: Exploratory hydrograph analysis. Hydrological Processes, 22(23), 4497-4506.
20. Schilling, K. E., Jindal, P., Basu, N. B., & Helmers, M. J. (2012). Impact of artificial subsurface drainage on groundwater travel times and baseflow discharge in an agricultural watershed, Iowa (USA). Hydrological Processes.
21. Skaggs, R. W., & Schilfgaarde, J. (1999). Drainage simulation models. Agricultural drainage., 469-500.
22. Stillman, J. S., Haws, N. W., Govindaraju, R., & Suresh C Rao, P. (2006). A semi-analytical model for transient flow to a subsurface tile drain. Journal of Hydrology, 317(1), 49-62.
23. Sun, J. (2012). Exploring edge complexity in remote-sensing vegetation index imageries. Journal of Land Use Science, 1-13.
24. Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resources Research, 33(2), 309-319.
25. Ducks Unlimited. (2010). Updating the National Wetlands Inventory (NWI) for Indiana (pp. 112).
26. USDA. (1992). Census of Agriculture United States Summary and State Data.
27. USDA. (2009). Census of Agriculture United States Summary and State Data.
28. Varner, B., Gress, T., White, S., & Robert, P. (2003). The effectiveness and economic feasibility of image-based agricultural tile maps. Paper presented at the Proceedings of the 6th International Conference on Precision Agriculture and Other Precision Resources Management, Minneapolis, MN, USA, 14-17 July, 2002.
29. Venkatesh, S., & Owens, R. (1990). On the classification of image features. Pattern Recognition Letters, 11(5), 339-349.
30. Verma, A., Cooke, R., & Wendte, L. (1996). Mapping subsurface drainage systems with color infrared aerial photographs. Urbana-Champaign: Department of Agricultural Engineering, University of Illinois.
31. Walter, M. F., Black,R.D. and Zwerman,P.J. (1979). Tile Flow Response in a Layered Soil. Trans,ASAE.
32. Wright, J., & Sands, G. (2001). Planning an Agricultural Subsurface Drainage System. College of Agricultural, Food and Environmental Sciences, University of Minnesota. BU-07685, 1-10.

CURRICULUM VITAE

Samapriya Roy

Academic Background

Indiana University, Indianapolis

Graduate student (M.S) Geology July 2013

Thesis: Remote Sensing & GIS applications for drainage detection and modeling in agricultural Watersheds

Visvesvaraya National Institute Of Technology (VNIT), Nagpur, India

Bachelor of Technology in Civil Engineering July 2010

Thesis: Social Network Analysis for Mapping Segmented Growth in Urban Cities in India.

Past Involvements

- Involved with USDA/NRCS project for Creating and Automating Potential Polygon Location and Site Characterization for Upland Storage in Watersheds, using Arc Map. A CTI (Compound Topographic Index) based Toolbox was developed using Model Builder (® ESRI) as a part of project deliverables to automate location of these potential sites.
(September 2011-August 2012)
- Teaching Assistant: Environmental Geology Course, G117 Spring 2012, G117 Fall 2012, G117 Spring 2013

Relevant Experience

- Title: Research Associate, Water Resource Management Group, Ganga River Basin Management Plan (GRBMP), under Ministry of Environment and Forests
Responsibilities: Included assessment of the hydrology of the Upper Ganga Basin and working on flow simulation and scenario development for flow calculation.
Organization: Indian Institute of Technology (IIT), Kanpur, January to June 2011.
- Title: Research Intern, Core Geospatial and Utilities(CGO) Business Unit
Responsibilities: Created a characteristic model called SMCPM (Scenario Modeling with Catchment Priority Model) which utilizes the SCN method accompanied with SMART and MAUT utilization for user based criterion input and priority output.
Organization: RMSI, Dehradun, 22nd June to 21st August 2010

Academic Internships

- “Pareto Optimal input parameterization for Water Quality Models using AHP”
Organization: Indian Institute of Science (IISc), Bangalore, June 2009
- “Water Quality Modeling for waste loading in natural streams applying QUAL2Kw”
Organization: Indian Institute of Science (IISc), Bangalore, December 2008
- “Numerical modeling of 1D Dam Break flow using Mac Cormack’s Scheme.”
Organization: Indian Institute of Technology (IIT) Kanpur, May to June 2008

Patents

Cyclical Hierarchical Modeling For Water Quality Model Based DSS Module In An Urban River System, Application Number- 1621/MUM/2011 A, International classification: G06F17/00, Early Publication, The Patent Office Journal 08/07/2011.

Scientific Peer-Reviewed Journal Publication

Roy, Samapriya, and Katpatal, Y.B (2011) Cyclical Hierarchical Modeling for Water Quality Model based DSS Module in an urban river system, Journal of Environmental Engineering, ASCE. Vol. 137, Number 12, 1176-1184.

International Conference Publications

Roy, Samapriya, and Katpatal, Y.B (2011). Non Transitive Modeling for Generating Hierarchical Model for an urban river system in India. Fourth International Perspective on Current & Future State of Water Resources & the Environment, EWRI-ASCE at National University Of Singapore, Singapore. (Presented by Roy, Samapriya.)

Conference Presentations And Publications (India)

- Roy, Samapriya, and Katpatal, Y. B (2010) Status Monitoring of Nag River in Nagpur Urban Area in Central India with relation to Waste Water Management. Third International Perspective on Current & Future State of Water Resources & the Environment, EWRI-ASCE at Indian Institute of Technology Madras.
- Roy, Samapriya, and Katpatal, Y. B (2010) Urban Patch and Segmented Growth Analysis in Nagpur urban city, modeling system dynamics using GIS. National Conference on “Sustainable Development of Urban Infrastructure” at VNIT Nagpur.
- Roy, Samapriya(2010) Input parameterization for Multi Criterion Decision Making in a Water Quality Model using Analytic Hierarchy Process. Indian Conference for Academic Research by Undergraduate Students at IIT Kanpur.