MAPPING METHODS METADATA FOR RESEARCH DATA

Tiffany C. Chao

International Digital Curation Conference
10 February 2015

GRADUATE SCHOOL OF LIBRARY AND INFORMATION SCIENCE
The iSchool at Illinois

Outline

- Background: Metadata generation for research data
- Introducing *methods metadata*
- Study design
- Preliminary findings
- Future directions

Generating metadata for research data

- Scientists have limited time/energy/resources to dedicate toward metadata generation
- Problematic for 'long-tail' science researchers who have heterogeneous data and insufficient support (Heidorn, 2008; Cragin et al., 2010)
- Increased requirements for public access to research data generated by government funding necessitate metadata provision (i.e. OSTP memo, Holdren, 2013)

Methods matter for data sharing and reuse

- Significance of research methods description
  - Documenting modifications to research protocol vital to integrity of dataset in field-based research
    (Karasti, Baker & Halkola, 2006; Karasti & Baker, 2008)
  - For trust and selection of data for reuse
    (Zimmerman, 2008; Van House, Butler & Schiff, 1998)

Methods metadata

- Focus on *methods metadata* in research

  the type of information needed for basic comprehension of how data were produced in the scientific research context

  - Highlights the importance of research methods description for describing research data for reuse

Methods metadata source?

- Limited use of automated approaches (i.e. workflow technologies) to record data production processes

- Role of journal publications as source for methods description
  - experimental procedure details in articles contribute to reuse decisions (Faniel & Jacobsen, 2010)
  - articles traditional mode for scientific communication with increased availability due to open access publishing (Brown, 2010)

Proposed inquiry

How can journal article content be utilized to generate metadata on data production methods for datasets?

- What metadata elements for methods from existing schemes map to journal article content?
- What gaps exist for methods metadata generation from journal article content?

Study Design

- Phase 1 – Mapping metadata for research methods
  - National Environmental Methods Index (NEMI) documentation
  - Sample of (24) peer-reviewed journal research articles from soil ecology

- Phase 2 – Extending mapping with existing metadata schemes for data
  - Ecological Metadata Language (EML)
  - Content Standard for Digital Geospatial Metadata (CSDGM)

NEMI metadata for methods

mandatory elements	Method Descriptive Name	Brief Method Summary
	Source Citation	Method Official Name
	Method type/ subcategory (pre-defined list)	Method Number/Identifier
	Media Name	Instrumentation
	Method Source	
optional elements	Scope and Application	Detection Limit Type
	Applicable Concentration	Interferences
	Max Holding Time	Sample Prep Methods
	Range	Concentration Range Units
	QC Requirements	Sample Handling
	Detection Limit Note	Precision Descriptor Notes

Methods metadata mapping with NEMI

APPLIED SOIL ECOLOGY 37 (2007) 161–168

available at www.sciencedirect.com

ScienceDirect

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

Earthworm population density and diversity in different-aged urban systems

K.M. Smetak[a,1], J.L. Johnson-Maynard[a,*], J.E. Lloyd[b,2]

[a] Soil and Land Resources, University of Idaho, Moscow, ID 83843-2339, United States

[b] The Urban Forestry Institute, 2239 Edgewood Ave. S., St. Louis Park, MN 55426, United States

ARTICLE INFO

Article history:
Received 16 November 2006
Received in revised form 19 May 2007
Accepted 1 June 2007

Keywords:
Earthworm density
Earthworm diversity
Urban soils
Landscape age

ABSTRACT

Earthworms are known to play important roles in soil profile development, nutrient cycling, and plant productivity where their population densities are high. Our objective was to understand the impact of disturbance and management intensity within different-aged urban landscapes on earthworm population density and diversity. Our study was conducted in Moscow, ID, USA. Earthworms and soils were collected from open turf in three residential yards less than 10 years old, three residential yards greater than 75 years old, and in three urban parks greater than 75 years old. Mean earthworm density was significantly different among urban park (437 individuals m^{-2}), old residential (121 individuals m^{-2}), and young residential sites (26 individuals m^{-2}). Bulk density of the 0–10 cm depth was significantly higher in young residential sites (1.59 g cm^{-3}) than in urban park (1.30 g cm^{-3}) or old residential sites (1.30 g cm^{-3}). Mean total soil C in the first 30 cm of soil was significantly different among the three landscape types (3.6 kg C m^{-2} in urban parks, 2.9 kg C m^{-2} in old residential, and 1.4 kg C m^{-2} in young residential sites). Total soil N content followed the same trend as soil C across landscape types. Fertilizer additions in urban park sites were associated with increases in earthworm density and total soil C and N content. Overall, high bulk density due to soil compaction in young landscapes may directly and indirectly limit earthworm density. Bulk density, however, tends to decrease while soil C and N content increase as urban systems mature and these changes are associated with increases in earthworm population density.

© 2007 Published by Elsevier B.V.

1. Introduction

Earthworm density and activity are influenced by factors such as food quality and quantity (Lee, 1985; Edwards and Bohlen, 1996; Curry, 2004), soil temperature and moisture (Berry and Jordan, 2001; Wever et al., 2001) and soil structure and texture (Nuutinen et al., 1998; Baker and Whitby, 2003). The relationship of earthworm activity to soil physical and chemical properties has been documented in forest (e.g. Whalen, 2004; Marhan and Scheu, 2005; Ammer et al., 2006), pasture (e.g. Baker et al., 1992; Decaëns et al., 2004; Winsome et al., 2006) and agricultural systems (e.g. Edwards et al., 1995; Lamandé et al., 2003), but has not been well studied in urban settings. Urban soils differ from wildland and agricultural soils in the type and degree of human

* Corresponding author. Tel.: +1 208 885 9245; fax: +1 208 885 7760.
E-mail addresses: smet4265@uidaho.edu (K.M. Smetak), jmaynard@uidaho.edu (J.L. Johnson-Maynard), JLloyd@RainbowTreecare.com (J.E. Lloyd).
[1] Tel.: +1 208 885 7505; fax: +1 208 885 7760.
[2] Tel.: +1 952 922 3810; fax: +1 952 252 4720.
0929-1393/$ – see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.apsoil.2007.06.004

APPLIED SOIL ECOLOGY 37 (2007) 161–168 163

Table 1 – Management practices for urban park (UP), young residential (YR), and old residential (OR) sites

Site	Age	Fertilizer	Fertilizer application frequency	Irrigation	Aeration[a]
UP 1	100	1.97 kg 100 m^{-2} of 25–5–10 fertilizer	Once/year	Every 10 days; June–September	N/A
UP 2	100	1.97 kg 100 m^{-2} of 21–3–14 fertilizer	Once/year	2.54 cm weekly; June–September	Once/year
UP 3	100	1.97 kg 100 m^{-2} of 21–3–14 fertilizer	Twice/year	2.54 cm weekly; June–September	Once/year
YR 1	8	1.8 kg 100 m^{-2} of 27–4–6 fertilizer[b]	Twice/year	Once nightly with sprinkler; June–September	N/A
YR 2	10	1.8 kg 100 m^{-2} of 27–4–6 fertilizer[b]	Once/year	Two to three times weekly with sprinkler; June–September	Once/year
YR 3	5	1.8 kg 100 m^{-2} of 27–4–6 fertilizer[b]	Once/year	Once daily with sprinkler; June–September	N/A
OR 1	75	N/A	N/A	N/A	N/A
OR 2	75	N/A	N/A	Once weekly with sprinkler; June–September	N/A
OR 3	95	N/A	N/A	<Once weekly with hose; June–August	N/A

N/A, Not applicable (management practice was not implemented).
[a] Aeration is the physical removal of multiple soil cores (2–2.5 cm diameter to a depth of between 3 and 8 cm) over the area of the lawn to improve the movement of air, water, and nutrients to the grass roots. Soil cores were removed with a mechanical aerator and left on the ground at all aerated sites.
[b] Homeowners reported adding locally available nitrogen-rich lawn fertilizer at the rate recommended on the label. Values are estimated using the N–P–K ratio of the typical lawn fertilizer available in the area and assuming a nitrogen application rate of 0.5 kg N 100 m^{-2} (the recommended application rate on the label). At all sites grass clippings were left on the surface to decompose.

2. Materials and methods

2.1. Site description

Study sites were located in Moscow, ID, USA (46°44′N, 116°48′W) and consisted of three urban parks (UP 1–3) greater than 75 years old, three young residential lawns (YR 1–3) less than 10 years old, and three old residential lawns (OR 1–3) greater than 7... years old. Site age was determined through interviews wit... homeowners and urban park managers as well as historic... markers in the case of old residential landscapes and urba... parks. Recently developed residential sites were generall... inhabited by the original owners who knew the exact ages o... their landscapes. The maximum distance between sites wa... 4.8 km which minimized microclimatic differences. Sites wher... insecticides were applied were not utilized due to possibl... negative impacts on earthworms.

Soils at all sites were formed in loess and had silt loam t... silty clay loam textures with depth. Litter layer thickness...

2.2. Field methods

Earthworm sampling was conducted during a 2-week period beginning at the end of May in 2004 and 2005. Each site was sampled once per year. Soil sampling was conducted during the 2004 earthworm sampling period. At each replicate site, four randomly located pits (15 cm × 25 cm) were dug to a depth of 30 cm for earthworm and soil collection. Earthworms were separated from soil by hand-sorting, placed in plastic Petri dishes lined with moist paper towels and taken back to the laboratory where they were counted. Earthworms were left in the Petri dishes for 48 h to void their guts of soil and were then weighed to calculate freshweight biomass. Earthworm density and biomass to the 30 cm depth were expressed on an area basis by dividing each value by the area of the sampling pit (375 cm^2). Earthworms were separated into three age classes – juvenile, adolescent, and mature – based on the development of the clitellum, the reproductive gland used for cocoon production that in mature earthworms generally forms an obvious band around the midsection segments. Juveniles had no obvious clitellum, adolescents had lightening of segments where the clitellum was developing, and mature earthworms had a fully developed clitellum. Mature earthworms were euthanized with ethanol and identified to the species level using Schwert's (1990) taxonomic key. Since the clitellum is only fully expressed in mature earthworms and is the main external feature used in earthworm identification, juveniles and adolescent earthworms were not identified. Bulk soil within each pit was collected from the 0–10, 10–20, and 20–30 cm depths for chemical analysis. Three replicate soil cores (131.9 cm^3) were taken from undisturbed soil adjacent to each pit from each depth for determination of bulk density (Grossman and Reinsch, 2002). Three replicate soil moisture cans were filled with soil from each depth within each pit for the determination of gravimetric soil moisture.

2.3. Laboratory analyses

Moisture tins containing field-moist soil were weighed, dried for 24 h at 105 °C, and reweighed to calculate gravimetric soil moisture content. Bulk density was calculated by dividing the oven-dried soil weight from each core by the core volume. Soil for chemical analysis was air-dried, gently crushed with a mortar and pestle, and passed through a 2 mm sieve prior to analysis. Soil texture was determined by the pipette method (Gee and Or, 2002). Total C and N were measured by dry combustion (Nelson and Sommers, 1982) in an Elementar Vario Max CNS analyzer (Elementar Analysensysteme, Hanau, Germany). Ammonium (NH_4^+-N) and nitrate (NO_3^--N) in 2 M KCl extracts (Keeney and Nelson, 1982) were measured using an automated colorimetric analyzer (Lachat Instruments, Milwaukee, WI, USA). Soil pH was determined on a 1:1 (soil:water) mixture.

2.4. Statistical analyses

All statistical analyses were conducted using SAS version 9.1 (SAS Inc., Cary, NC, USA). Data were analyzed using ANOVA techniques (PROC GLM) and mean separation carried out with Fisher's least significant difference. Response variables included earthworm density, bulk density, soil moisture, pH, total carbon, total nitrogen, and inorganic nitrogen and treatment classifications were urban park, old residential, and young residential sites. Significant differences were determined assuming a 95% confidence level. Earthworm density, earthworm weight, and soil inorganic nitrogen values had non-normal distributions, and were therefore analyzed using log-transformed data. Additionally, correlation analysis was performed to examine the relationship of earthworm density to bulk density, soil moisture, and total carbon and nitrogen, in urban park, old residential, and young residential sites.

3. Results and discussion

3.1. Earthworm density

Earthworm density was significantly different among all landscape types ($p < 0.0001$). The highest densities were observed in urban park sites (437 individuals m^{-2}), followed by old residential sites (121 individuals m^{-2}), with the lowest densities observed in young residential sites (26 individuals m^{-2}) (Fig. 1). Earthworm freshweight in urban park sites (...12 g m^{-2}) was significantly higher ($p < 0.0001$) than in old residential (28.08 g m^{-2}) or young residential (4.69 g m^{-2}) sites (Fig. 2).

Low earthworm density in the young residential sites is likely due to a combination of lack of adequate time for earthworm invasion from surrounding areas and the soil's inability to provide proper habitat. Documented rates of spread for introduced earthworms range from 4 to 14 m year^{-1} (Edwards and Bohlen, 1996). Depending on the distance of new urban landscapes to sources of earthworms such as forest and agricultural fields, establishment of populations from immigration in a 10-year period may not be likely. A more important source of earthworms, in these systems, would be through

Method Citation

Instrumentation

Method Descriptive Name

Brief Method Summary

Media Name

Article example: Smetak, K. M., Johnson-Maynard, J. L., & Lloyd, J. E. (2007). Earthworm population density and diversity in different-aged urban systems. Applied Soil Ecology, 37(1), 161–168.

NEMI

CSDGM – “Lineage”

[mandatory]

Process Step
- Description
- Date
- Time
- Contact
- Source Produced/Used Citation Abbreviation

[mandatory if applicable]

Source
- Information
- Citation/Abbreviation
- Time Period of Content
- Contribution
- Type

NEMI

- Descriptive Name
- Source Citation
- Method type/ subcategory
- Media Name
- Brief Method Summary
- Official Name
- Method #
- Instrumentation
- Method Source
- Scope and Application
- Applicable Concentration
- Max Holding Time
- Range
- QC Requirements
- Detection Limit Note
- Precision Descriptor Notes
- Detection Limit Type
- Interferences
- Sample Prep Methods
- Concentration Range Units
- Sample Handling

EML – “methods module”

[mandatory]

methodsType
methodsStep
procedureStepType
- description
- citation
- protocol
- instrumentation
- software
- substep

[mandatory if applicable]

dataSource
sampling
- studyExtent
- spatialSamplingUnits
- description

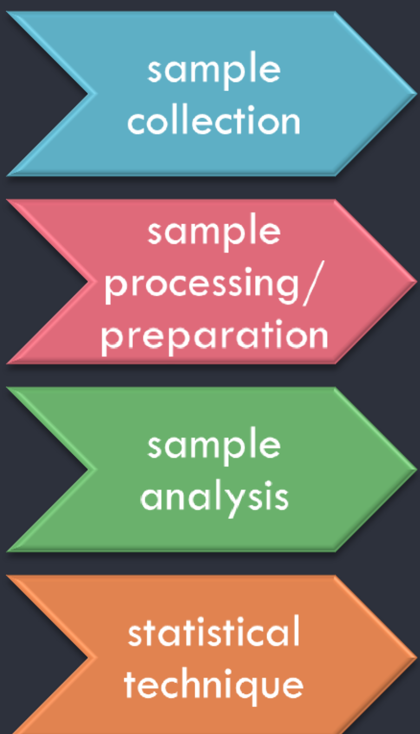
qualityControl
- Description, citation, protocol, instrumentation, software, substep

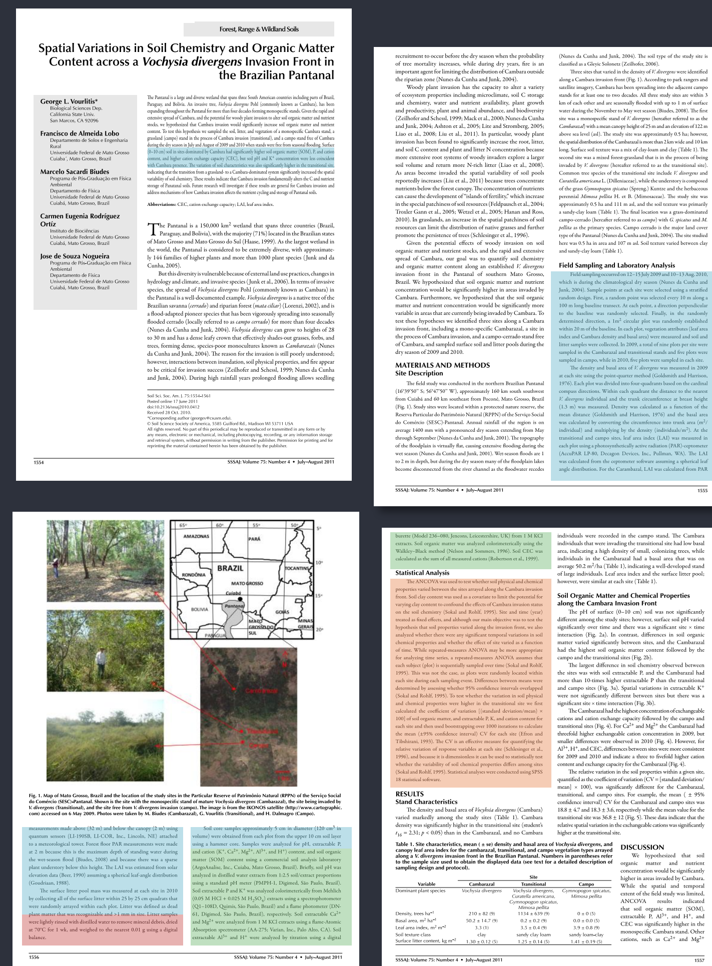
Preliminary findings & Discussion

- Methods information from journal articles amenable to metadata generation
- Common methods-related elements identified across metadata schemes
- Gaps in both metadata schemes and journal article content revealed for methods description

Journal article structure and mapping

NEMI method types include:




Forest, Range & Wildland Soils

Spatial Variations in Soil Chemistry and Organic Matter Content across a *Vochysia divergens* Invasion Front in the Brazilian Pantanal

- "Methods and Materials" section most robust area for methods description
- Descriptions of "processing" and "analysis" not as explicit in article content

Article example: Vourlitis, G. L. et al. (2011). Spatial Variations in Soil Chemistry and Organic Matter Content across a Vochysia divergens Invasion Front in the Brazilian Pantanal. Soil Sci. Soc. Am. J., 75(4), 1554–1561.

Methods metadata mapping: connecting data and processes

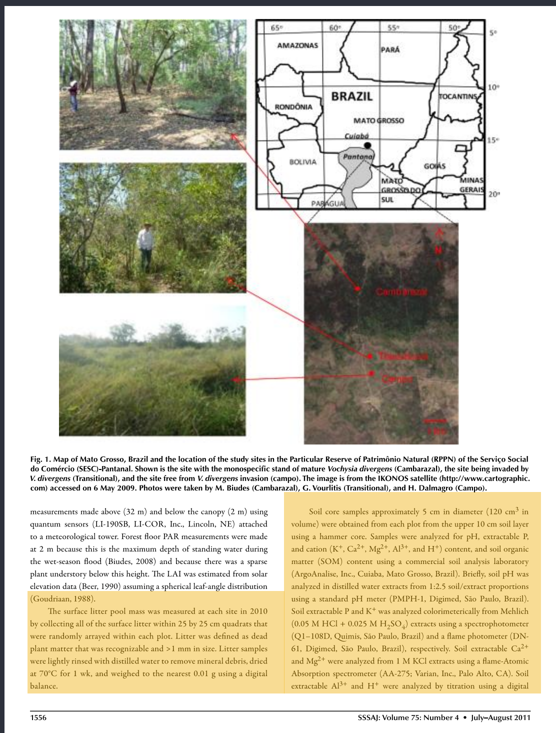


Fig. 1. Map of Mato Grosso, Brazil and the location of the study sites in the Particular Reserve of Patrimônio Natural (RPPN) of the Serviço Social do Comércio (SESC)-Pantanal. Shown is the site with the monospecific stand of mature *Vochysia divergens* (Cambarazal), the site being invaded by *V. divergens* (Transitional), and the site free from *V. divergens* invasion (campo). The image is from the IKONOS satellite (http://www.cartographic.com) accessed on 6 May 2009. Photos were taken by M. Biudes (Cambarazal), G. Vourlitis (Transitional), and H. Dalmagro (Campo).

measurements made above (32 m) and below the canopy (2 m) using quantum sensors (LI-190SB, LI-COR, Inc., Lincoln, NE) attached to a meteorological tower. Forest floor PAR measurements were made at 2 m because this is the maximum depth of standing water during the wet-season flood (Biudes, 2008) and because there was a sparse plant understory below this height. The LAI was estimated from solar elevation data (Beer, 1990) assuming a spherical leaf-angle distribution (Goudriaan, 1988).

The surface litter pool mass was measured at each site in 2010 by collecting all of the surface litter within 25 by 25 cm quadrats that were randomly arrayed within each plot. Litter was defined as dead plant matter that was recognizable and >1 mm in size. Litter samples were lightly rinsed with distilled water to remove mineral debris, dried at 70°C for 1 wk, and weighed to the nearest 0.01 g using a digital balance.

Soil core samples approximately 5 cm in diameter (120 cm^3 in volume) were obtained from each plot from the upper 10 cm soil layer using a hammer core. Samples were analyzed for pH, extractable P, and cation (K^+, Ca^{2+}, Mg^{2+}, Al^{3+}, and H^+) content, and soil organic matter (SOM) content using a commercial soil analysis laboratory (ArgoAnalise, Inc., Cuiaba, Mato Grosso, Brazil). Briefly, soil pH was analyzed in distilled water extracts from 1:2.5 soil/extract proportions using a standard pH meter (PMPH-1, Digimed, São Paulo, Brazil). Soil extractable P and K^+ was analyzed colorimeterically from Mehlich (0.05 M HCl + 0.025 M H_2SO_4) extracts using a spectrophotometer (Q1–108D, Quimis, São Paulo, Brazil) and a flame photometer (DN-61, Digimed, São Paulo, Brazil), respectively. Soil extractable Ca^{2+} and Mg^{2+} were analyzed from 1 M KCl extracts using a flame-Atomic Absorption spectrometer (AA-275; Varian, Inc., Palo Alto, CA). Soil extractable Al^{3+} and H^+ were analyzed by titration using a digital

1556 SSSAJ: Volume 75: Number 4 • July–August 2011

(Vourlitis, G. L. et al., 2011)

Method Identifier	**VALBRSN-2011-01**
Method Descriptive Name	Surface litter processing
Brief Method Summary	The surface litter pool mass was measured at each site in 2010 by collecting all of the surface litter within 25 by 25 cm quadrats that were randomly arrayed within each plot. Litter samples were lightly rinsed with distilled water to remove mineral debris, dried at 70°C for 1 wk, and weighed to the nearest 0.01 g using a digital balance.
Media type	Surface litter
Method Identifier	**VALBRSN-2011-02**
Method Descriptive Name	Soil core sample collection & analysis
Brief Method Summary	Soil core samples approximately 5 cm in diameter (120 cm3 in volume) were obtained from each plot from the upper 10 cm soil layer using a hammer core. Samples were analyzed for pH, extractable P, and cation content, and soil organic matter (SOM) content using a commercial soil analysis laboratory. Briefly, soil pH was analyzed in distilled water extracts from 1:2.5 soil/extract proportions using a standard pH meter
Media type	Soil core samples
Method source	Commercial soil analysis laboratory (ArgoAnalise, Inc., Cuiaba, Mato Grosso, Brazil) used for pH, extractable P, and cation content, and soil organic matter (SOM)
Instrumentation	pH meter for soil (PMPH-1, Digimed, São Paulo, Brazil).

Methods metadata map: extending to other metadata schemes

Method Identifier	VALBRSN-2011-01
Method Descriptive Name	Surface litter processing
Brief Method Summary	The surface litter pool mass was measured at each site in 2010 by collecting all of the surface litter within 25 by 25 cm quadrats that were randomly arrayed within each plot. Litter samples were lightly rinsed with distilled water to remove mineral debris, dried at 70°C for 1 wk, and weighed to the nearest 0.01 g using a digital balance.
Media type	Surface litter

CSDGM	
<Process Step> Description	The surface litter pool mass was measured at each site in 2010 by collecting all of the surface litter within 25 by 25 cm quadrats that were randomly arrayed within each plot.
Source Produced Citation	Surface litter VALBRSN-2011-A
<Process Step> Description	Litter samples were lightly rinsed with distilled water to remove mineral debris, dried at 70°C for 1 wk, and weighed to the nearest 0.01 g using a digital balance.
Source Used Citation	VALBRSN-2011-A
Source Produced Citation	Weight (g) of surface litter VALBRSN-2011-b

Method Identifier	VALBRSN-2011-02
Method Descriptive Name	Soil core sample collection & analysis
Brief Method Summary	Soil core samples approximately 5 cm in diameter (120 cm3 in volume) were obtained from each plot from the upper 10 cm soil layer using a hammer core. Samples were analyzed for pH, extractable P, and cation content, and soil organic matter (SOM) content using a commercial soil analysis laboratory. Briefly, soil pH was analyzed in distilled water extracts from 1:2.5 soil/extract proportions using a standard pH meter
Media type	Soil core samples
Method source	Commercial soil analysis laboratory (ArgoAnalise, Inc., Cuiaba, Mato Grosso, Brazil) used for pH, extractable P, and cation content, and soil organic matter (SOM)
Instrumentation	pH meter for soil (PMPH-1, Digimed, São Paulo, Brazil)

EML	
dataSource	Soil core samples
Procedure description	Soil core samples approximately 5 cm in diameter (120 cm3 in volume) were obtained from each plot from the upper 10 cm soil layer using a hammer core.
Procedure description	Samples were analyzed for pH, extractable P, and cation content, and soil organic matter (SOM) content using a commercial soil analysis laboratory.
subStep	Briefly, soil pH was analyzed in distilled water extracts from 1:2.5 soil/extract proportions using a standard pH meter
instrumentation	pH meter for soil (PMPH-1, Digimed, São Paulo, Brazil)

Metadata schemes: common methods elements

Prevalent methods-related elements available in journal article content:

"description" "citation" "sampling"

- Common elements an indicator of essential information for methods metadata

Metadata scheme gaps for methods

- Study site information consistent across journal articles

recruitment to occur before the dry season when the probability of tree mortality increases, while during dry years, fire is an important agent for limiting the distribution of Cambara outside the riparian zone (Nunes da Cunha and Junk, 2004).

Woody plant invasion has the capacity to alter a variety of ecosystem properties including microclimate, soil C storage and chemistry, water and nutrient availability, plant growth and productivity, plant and animal abundance, and biodiversity (Zeilhofer and Schessl, 1999; Mack et al., 2000; Nunes da Cunha and Junk, 2004; Ashton et al., 2005; Lite and Stromberg, 2005; Liao et al., 2008; Liu et al., 2011). In particular, woody plant invasion has been found to significantly increase the root, litter, and soil C content and plant and litter N concentration because more extensive root systems of woody invaders explore a larger soil volume and return more N-rich litter (Liao et al., 2008). As areas become invaded the spatial variability of soil pools reportedly increases (Liu et al., 2011) because trees concentrate nutrients below the forest canopy. The concentration of nutrients can cause the development of "islands of fertility," which increase in the special patchiness of soil resources (Feldpausch et al., 2004; Troxler Gann et al., 2005; Wetzel et al., 2005; Hanan and Ross, 2010). In grasslands, an increase in the spatial patchiness of soil resources can limit the distribution of native grasses and further promote the persistence of trees (Schlesinger et al., 1996).

Given the potential effects of woody invasion on soil organic matter and nutrient stocks, and the rapid and extensive spread of Cambara, our goal was to quantify soil chemistry and organic matter content along an established *V. divergens* invasion front in the Pantanal of southern Mato Grosso, Brazil. We hypothesized that soil organic matter and nutrient concentration would be significantly higher in areas invaded by Cambara. Furthermore, we hypothesized that the soil organic matter and nutrient concentration would be significantly more variable in areas that are currently being invaded by Cambara. To test these hypotheses we identified three sites along a Cambara invasion front, including a mono-specific Cambarazal, a site in the process of Cambara invasion, and a campo-cerrado stand free of Cambara, and sampled surface soil and litter pools during the dry season of 2009 and 2010.

MATERIALS AND METHODS

Site Description

The field study was conducted in the northern Brazilian Pantanal (16°39'50" S; 56°47'50" W), approximately 160 km south southwest from Cuiabá and 60 km southeast from Poconé, Mato Grosso, Brazil (Fig. 1). Study sites were located within a protected nature reserve, the Reserva Particular do Patrimônio Natural (RPPN) of the Serviço Social do Comércio (SESC)-Pantanal. Annual rainfall of the region is on average 1400 mm with a pronounced dry season extending from May through September (Nunes da Cunha and Junk, 2001). The topography of the floodplain is virtually flat, causing extensive flooding during the wet season (Nunes da Cunha and Junk, 2001). Wet-season floods are 1 to 2 m in depth, but during the dry season many of the floodplain lakes become disconnected from the river channel as the floodwater recedes (Nunes da Cunha and Junk, 2004). The soil type of the study site is classified as a Gleyic Solonetz (Zeilhofer, 2006).

Three sites that varied in the density of *V. divergens* were identified along a Cambara invasion front (Fig. 1). According to park rangers and satellite imagery, Cambara has been spreading into the adjacent campo stands for at least one to two decades. All three study sites are within 3 km of each other and are seasonally flooded with up to 1 m of surface water during the November to May wet season (Biudes, 2008). The first site was a monospecific stand of *V. divergens* (hereafter referred to as the *Cambarazal*) with a mean canopy height of 25 m and an elevation of 122 m above sea level (asl). The study site was approximately 0.5 ha; however, the spatial distribution of the Cambarazal is more than 2 km wide and 10 km long. Surface soil texture was a mix of clay-loam and clay (Table 1). The second site was a mixed forest-grassland that is in the process of being invaded by *V. divergens* (hereafter referred to as the transitional site). Common tree species of the transitional site include *V. divergens* and *Curatella americana* L. (Dilleniaceae), while the understory is composed of the grass *Gymnopogon spicatus* (Spreng.) Kuntze and the herbaceous perennial *Mimosa pellita* H. et B. (Mimosaceae). The study site was approximately 0.5 ha and 111 m asl, and the soil texture was primarily a sandy-clay loam (Table 1). The final location was a grass-dominated campo-cerrado (hereafter referred to as *campo*) with *G. spicatus* and *M. pellita* as the primary species. Campo cerrado is the major land cover type of the Pantanal (Nunes da Cunha and Junk, 2004). The site studied here was 0.5 ha in area and 107 m asl. Soil texture varied between clay and sandy-clay loam (Table 1).

Field Sampling and Laboratory Analysis

Field sampling occurred on 12–15 July 2009 and 10–13 Aug. 2010, which is during the climatological dry season (Nunes da Cunha and Junk, 2004). Sample points at each site were selected using a stratified random design. First, a random point was selected every 10 m along a 100 m long baseline transect. At each point, a direction perpendicular to the baseline was randomly selected. Finally, in the randomly determined direction, a $1 m^2$ circular plot was randomly established within 20 m of the baseline. In each plot, vegetation attributes (leaf area index and Cambara density and basal area) were measured and soil and litter samples were collected. In 2009, a total of nine plots per site were sampled in the Cambarazal and transitional stands and five plots were sampled in campo, while in 2010, five plots were sampled in each site.

The density and basal area of *V. divergens* was measured in 2009 at each site using the point-quarter method (Goldsmith and Harrison, 1976). Each plot was divided into four-quadrants based on the cardinal compass directions. Within each quadrant the distance to the nearest *V. divergens* individual and the trunk circumference at breast height (1.3 m) was measured. Density was calculated as a function of the mean distance (Goldsmith and Harrison, 1976) and the basal area was calculated by converting the circumference into trunk area (m^2/individual) and multiplying by the density (individuals/m^2). At the transitional and campo sites, leaf area index (LAI) was measured in each plot using a photosynthetically active radiation (PAR)-ceptometer (AccuPAR LP-80, Decagon Devices, Inc., Pullman, WA). The LAI was calculated from the ceptometer software assuming a spherical leaf angle distribution. For the Carambazal, LAI was calculated from PAR

SSSAJ: Volume 75: Number 4 • July–August 2011 1555

Fig. 1. Map of Mato Grosso, Brazil and the location of the study sites in the Particular Reserve of Patrimônio Natural (RPPN) of the Serviço Social do Comércio (SESC)-Pantanal. Shown is the site with the monospecific stand of mature *Vochysia divergens* (Cambarazal), the site being invaded by *V. divergens* (Transitional), and the site free from *V. divergens* invasion (campo). The image is from the IKONOS satellite (http://www.cartographic.com) accessed on 6 May 2009. Photos were taken by M. Biudes (Cambarazal), G. Vourlitis (Transitional), and H. Dalmagro (Campo).

(Vourlitis, G. L. et al., 2011)

Study site details include:

- Longitude & latitude
- Geographic location name
- Average precipitation, humidity
- Soil type identification
- Local vegetation
- Site history (i.e. natural disturbances)

Schemes primarily accommodate geographic coordinates but not context description.

High level of detail for study site may warrant alternative representation in metadata record (i.e. citation to primary article).

Element coverage in articles

Some methods-related metadata elements not easily discernable in journal article content:

- Data quality processes
  - Generally not an explicit section, processes may be embedded with other procedures
- Time and date precision
  - Evident for “data collecting” but not processing or analysis

Access to additional resources may be required to obtain and/or verify description for these elements

Summary

- Use of journal article content for methods metadata generation reveals:
  - Metadata schemes generally accommodate documenting the relationships between data sources and processes
  - Common elements across schemes to support methods description
  - Potential enhancements to existing metadata schemes
  - Some aspects of methods description may require additional information resources

Future Directions

- Compare content of existing data metadata records with affiliated journal articles to understand application of scheme
  - Preliminary analysis indicates direct use of journal article content for generating methods metadata

- Investigate more systematic approach for identifying and documenting methods information from journal articles
  - Basis for potential for automation of methods metadata

- Extend journal article content to other areas of a metadata scheme beyond methods description

Conclusions

- Journal article content, as a whole, provide a robust source for methods metadata
  - Implications for journal publishers to maintain and support rich methods description in research articles
  - Other discipline research articles may vary in level of methods description

- Supporting data curation services
  - Understanding an unobtrusive approaches for data curation professionals to obtain metadata from researchers

Acknowledgements

- Dr. Carole Palmer, Dr. Michelle Wander, Dr. Jane Greenberg, and Dr. Catherine Blake
- Center for Informatics Research in Science and Scholarship
- Thomson Reuters Doctoral Dissertation Proposal award

Thank You!

Tiffany Chao

tchao@illinois.edu

References

- Brown, C. (2010). Communication in the sciences. Annual review of information science and technology, 44(1), 285–316.
- Faniel, I. M., & Jacobsen, T. E. (2010). Reusing scientific data: How earthquake engineering researchers assess the reusability of colleagues' data. Computer Supported Cooperative Work (CSCW), 19(3-4), 355–375. doi:10.1007/s10606-010-9117-8
- Heidorn, P. B. (2008). Shedding light on the dark data in the long tail of science. Library Trends, 57(2), 280–299.
- Holdren, J. P. (2013, February 22). Expanding public access to the results of federally funded research. Washington, DC: The White House. U.S. Executive Office of the President. Office of Science and Technology Policy. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp ostp_public_access_memo_2013.pdf
- Karasti, H., Baker, K. S., & Halkola, E. (2006). Enriching the notion of data curation in escience: Data managing and information infrastructure in the Long Term EcologicalResearch (LTER) network. Computer Supported Cooperative Work (CSCW), 15(4), 321-358. Springer. doi:10.1007/s10606-006-9023-2
- Karasti, H., & Baker, K. S. (2008). Digital data practices and the long term ecological research program growing global. The International Journal of Digital Curation, 2(3), 42 – 58. Retrieved from http://ijdc.net/index.php/ijdc/article/view/86
- Van House, N. A., Butler, M. H., & Schiff, L. R. (1998). Cooperative knowledge work and practices of trust: sharing environmental planning data sets. In Proceedings of the 1998 ACM conference on Computer supported cooperative work (pp. 335–343).
- Zimmerman, A. S. (2008). New knowledge from old data: The role of standards in the sharing and reuse of ecological data. Science, Technology & Human Values. doi:10.1177/0162243907306704