Selecting Color-based Tracers and Classifying Sediment Sources in the Assessment of Sediment Dynamics Using Sediment Source Fingerprinting

Louise R.M. Barthod, Kui Liu, David A. Lobb,* Philip N. Owens, Núria Martínez-Carreras, Alexander J. Koiter, Ellen L. Petticrew, Gregory K. McCullough, Cenwei Liu, and Leticia Gaspar

Abstract

The use of sediment color as a fingerprint property to determine sediment sources is an emerging technique that can provide a rapid and inexpensive means of investigating sediment sources. The present study aims to test the feasibility of color fingerprint properties to apportion sediment sources within the South Tobacco Creek Watershed (74 km²) in Manitoba, Canada. Suspended sediment from 2009 to 2011 at six monitoring stations and potential source samples along the main stem of the creek were collected. Reflectance spectra of sediments and source materials were quantified using a diffuse reflectance spectrometer, and 16 color coefficients were derived from several color space models. Canonical discriminant analysis was used to recategorize and downsize sediment source groups. After the linear additive test and stepwise discriminant function analysis, four color coefficients were chosen to fit the Stable Isotope Analysis in R model. Consistent with the conventional fingerprinting approach, the color fingerprint results demonstrated a switch in the dominant sediment source between the headwaters and the outlet of the watershed, with the main sources being topsoil in the upper reaches, whereas outcrop shale and stream bank materials dominated in the lower reaches. The color fingerprinting approach can be integrated with conventional fingerprints (e.g., geochemical and fallout radionuclide properties) to improve source discrimination, which is a key component for source ascription modeling. We concluded that the use of color fingerprints is a promising, cost-effective technique for sediment source fingerprinting.

Core Ideas

• Defining sources and testing the behavior of tracers are critical for sediment source apportionment.
• The predominant sources of river sediment varied at different reaches of the study creek.
• Integrating color and conventional fingerprinting techniques likely improves source apportionment.
• Color fingerprinting is a promising, cost-effective technique for sediment source ascription.

Copyright © 2015 American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America. 5585 Guilford Rd., Madison, WI 53711 USA. All rights reserved.

doi:10.2134/jeq2015.01.0043
Received 27 Jan. 2015.
Accepted 12 May 2015.
*Corresponding author (david.lobb@umanitoba.ca).

Agricultural nonpoint-source pollution can adversely affect surface water quality because inputs of nutrients such as phosphorus (P) and nitrogen (N) cause eutrophication, resulting in excessive plant growth, depletion of oxygen, and, as a consequence, changes in the abundance and diversity of organisms (Chambers et al., 2008). Over the past few decades, this phenomenon has been widely observed in Lake Winnipeg, Manitoba, Canada, the tenth largest freshwater lake in the world (23,750 km²) (Environment Canada, 2011). Indeed, lake water quality has significantly decreased since the mid-1990s, with a proliferation of algal blooms due to a recent and rapid increase in loading and concentration of P (Schindler et al., 2012). The Lake Winnipeg Basin drains ~1,000,000 km², with nearly 65 million ha of farmland (65%), including parts of four Canadian provinces (Alberta, Saskatchewan, Manitoba, and Ontario) and four US states (Montana, North Dakota, South Dakota, and Minnesota). Agriculture in Manitoba contributes approximately 5% of the total N and 15% of the total P loading into Lake Winnipeg (1994–2001) (Lake Winnipeg Stewardship Board, 2006). Therefore, agricultural beneficial management practices (BMPs) have been implemented to minimize nutrient loading in Lake Winnipeg. Understanding the processes involved, such as the interactions of sediment with nutrients, is required to be able to evaluate the effectiveness of the proposed BMPs. Against this background, identifying the sources of sediment within agricultural basins across the Canadian Prairies plays an important role for the conservation of Lake Winnipeg.

Published September 16, 2015

 Abbreviations: BMP, beneficial management practice; CIE, Commission Internationale de l’Eclairage; DFA, discriminant function analysis; FRN, fallout radionuclide; RI, Redness Index; SIAR, Stable Isotope Analysis in R; STCW, South Tobacco Creek Watershed.
because the sources of sediment are likely to influence its contaminant loading.

Fingerprinting techniques can be used to distinguish the sources of sediment and apportion their contributions within a watershed, and as such there has been a rapid growth in studies using fingerprinting to address a range of questions in watersheds throughout the world (Walling, 2005; Walling, 2013). A variety of radiochemical, biogeochemical, and physical properties have been successfully used worldwide in the application of environmental fingerprinting techniques, such as geochemistry and/or radiochemistry (e.g., Foster and Walling, 1994; Collins et al., 1997), mineral magnetism (e.g., Walden et al., 1997), particle size (e.g., Stone and Sauderson, 1992), and more recently soil color (e.g., Martinez-Carreras et al., 2010a; Martinez-Carreras et al., 2010b) and infrared spectroscopy (e.g., Poulenc et al., 2009; Evrard et al., 2013; Brosinsky et al., 2014b; Verheyen et al., 2014). However, very limited research using these techniques has been performed in the Lake Winnipeg Basin despite a pressing need for information on sediment and associated nutrient sources. A recent study (Koiter et al., 2013a) using geochemical and fallout radionuclide (FRN) fingerprints in the South Tobacco Creek Watershed (STCW), Manitoba, produced valuable insights about the development of this technique. In this watershed, a switch in the predominant source of sediment between the headwaters and the watershed outlet was observed, highlighting the importance of the sampling location in relation to the scale and geomorphic connectivity of the watershed. However, this technique, which uses geochemical and FRN fingerprint properties, is highly labor intensive and costly, and there is a need to develop rapid and inexpensive means of investigating sediment sources, especially from a watershed management perspective.

Color-based fingerprints used to quantify sediment sources successfully in small-to-medium scale catchments (e.g., Martinez-Carreras et al., 2010a; Martinez-Carreras et al., 2010b; Brosinsky et al., 2014a; Brosinsky et al., 2014b; Verheyen et al., 2014) or during runoff events in headwater catchments (e.g., Legout et al., 2013) could provide a rapid, inexpensive, and straightforward method to investigate sediment sources. Soil color is commonly described quantitatively using a Munsell soil color chart (Munsell Color Company, 1975), which is intuitively designed to reflect our perception of color and its variations (Viscarra Rossel et al., 2006). However, the values obtained using this approach do not represent a continuous physical variable and, as a consequence, cannot be used as a fingerprint property. Color is a three-dimensional phenomenon (Hunt and Pointer, 2011) and can be represented in color space models whereby individual colors are specified by points in these spaces. Those parameters are continuous physical variables that can be used to describe soil color and thus can be used to quantify the sources of sediment in the application of sediment source fingerprinting.

To a large extent, the color-based fingerprinting approach has failed to discriminate sediment sources unambiguously in medium- and large-sized catchments (Martinez-Carreras et al., 2010a) where heterogeneous pedology and geology as well as intrasource variability and source overlap are observed. Yet, this approach has been typically based on suspended sediment samples collected at the catchment outlet only and therefore does not fully address the influence of basin topography, geomorphology, and connectivity on sediment dynamics. However, it is now well known that only a fraction of sediment eroded within a catchment will reach the basin outlet (Fryirs, 2013). Considering the (dis)connectivity and storage that could occur throughout the watershed (Fryirs, 2013), it is important to fully understand the sediment dynamics and to determine the sediment sources at various scales. With the exception of Martinez-Carreras et al. (2010a) and Verheyen et al. (2014), there have been no attempts to test the color-based fingerprints at a variety of scales in medium- and large-sized catchments despite realization of the need for this (Brosinsky et al., 2014a). The present study explored the potential of color-based fingerprints as a means of apportioning sediment contributions at a variety of scales in STCW. The results are compared with the findings of the “conventional” geochemical and FRN fingerprinting approach undertaken in the same watershed (Koiter et al., 2013a) to assess the efficiency of color parameters as fingerprint properties, particularly their feasibility to detect changes in the predominant sources along the study creek.

Materials and Methods

Study Area

The STCW (74.4 km²) lies within the Lake Winnipeg Basin and is situated approximately 100 km southwest of Winnipeg (Fig. 1). The STCW has been the subject of several scientific projects exploring the effects of small dams, reservoirs, and agricultural BMPs on water and air quality, hydrology, runoff, soil erosion, and nutrient and sediment losses (e.g., Rawn et al., 1999; Li et al., 2007; Tiessen et al., 2010; Li et al., 2011; Tiessen et al., 2011; Liu et al., 2013).

Within the STCW, three physiographic regions can be distinguished from west to east: (i) the upper reaches (above 410 m), which lies in the undulating glacial tills; (ii) the Manitoba Escarpment (between 410 and 320 m), where the elevation drops nearly 90 m in less than 5 km; and (iii) the lower reaches (below 320 m), which lies in the lacustrine sediments of glacial Lake Agassiz (Fig. 2). Soils are primarily clay loams formed on moderately to strongly calcareous glacial till that overlay Cretaceous shale bedrock (Agriculture and Agri-Food Canada, 2013). The South Tobacco Creek flows across the Manitoba Escarpment descending over 200 m, incising the escarpment and exposing the underlying shale bedrock. Because the point of incision (the “knick point”) migrates upstream, it has produced steep valley walls and actively eroding cut-bank shale bedrock outcrops along the channel, some exceeding 20 m in height (Koiter et al., 2013a). Although the main outcrop is located in the middle of the watershed, there are also small outcrops in the headwaters. This shale is friable and can easily be broken down into fine-grained sediments under the effects of fluvial processes as well as freeze-thaw and wetting-drying cycles and as such potentially represents an important source of sediment (Koiter et al., 2013a). At Miami, the outlet of the STCW, the South Tobacco Creek drains into the Tobacco Creek, which flows north into Lake Winnipeg (Fig. 1).

Agriculture accounts for 74% of land use in the STCW. Forests cover 25% of the watershed and are mostly limited to...
valley walls that are unsuitable for cultivation. The remainder is mostly urban/transportation areas and surface water (1%) (Land Cover 2010 using LiDAR imagery) (Melvin, 2010).

The climate of the study area is classified as continental subhumid and has warm summers and cold winters (range, −30°C to +30°C). The mean annual precipitation (1971–2000, Environment Canada weather station, Deerwood) is estimated to be 566 mm, with 27% occurring as snowfall (151 cm). The average temperature for the same period is 3°C, with monthly average temperatures below 0°C from November to March (Environment Canada, 2013a).

Along the study stream, daily discharge was measured at two gauging stations operated by the Water Survey of Canada (WSC) (Environment Canada, 2013b). One station is below the escarpment (site 4, WSC 05OF023), and the other is at the STCW outlet (site 6, WSC 05OF017) (Fig. 1). The high discharges occur during the snowmelt (April) and major rainfall events (May–July) (Fig. 3). Twenty-six small dams and reservoirs were constructed between 1985 and 1996 in the headwaters to reduce peak flows and flooding (Agriculture and Agri-Food Canada, 2012). Moreover, significant natural damming occurs throughout the watershed due to beavers (*Castor canadensis*), which contributes to the modification of the hydrology of the watershed.

The STCW presents heterogeneous climatic, topographical, geomorphological, and hydrological characteristics that may influence the sediment dynamics throughout the watershed. To better appreciate this phenomenon, six sediment monitoring stations were established along the main stem of the creek (Fig. 1) to provide information on the sediment sources at various spatial scales.

![Fig. 1. Overview of the Tobacco Creek Watershed (Manitoba) and in-stream monitoring stations.](image1)

![Fig. 2. Longitudinal profile of the South Tobacco Creek and the Manitoba Escarpment across South Tobacco Creek Watershed (modified from Koiter et al., 2013a).](image2)
Fig. 3. Daily discharge at two sites in the South Tobacco Creek Watershed (site 4: south branch of the South Tobacco Creek near highway 240; site 6: south Tobacco Creek at Miami) (source data: Environment Canada, 2013b).

Collection and Preparation of Sediment and Source Samples

Samples of suspended sediment were collected at each in-stream monitoring station (Fig. 1) using a passive time-integrated sediment sampler following the design of Phillips et al. (2000). In brief, the samplers work on the principle that the reduced flow velocity of water entering the large chamber promotes settling of sediment (for details, see Phillips et al. [2000]). At sites 1, 2, 4, 5, and 6, suspended sediment samples were collected between 2009 and 2011; site 3 was added in 2011. The drainage areas for sites 1, 2, 3, 4, 5, and 6 are 44, 188, 1600, 3645, 6644, and 7441 ha, respectively. At each site, two samplers were installed during the ice-free period (Apr.–Nov.). The samplers were regularly inspected for maintenance, and the suspended sediment samples were collected two or three times per year, depending on the magnitude and duration of the major flow periods. The slurry of sediment and water collected from the samplers was allowed to settle, and then the water was decanted for sediment collection. The suspended sediments were air-dried and disaggregated by hand with a marble roller or gently with a mortar and pestle. For each site, the sediment samples from the two samplers were combined before analysis. Particle-size distributions of the physically disaggregated sediment were measured using laser diffraction (Mastersizer 2000, Malvern) in the range of 0.02 to 2000 μm. For each site, the grain-size composition of suspended-sediment samples was predominantly <63 μm. Of the 42 bulk sediment samples collected between 2009 and 2011, 27 were analyzed by spectrophotometry (Table 1). The remaining 15 samples was not analyzed due to inadequate dry mass, although other studies have collected small-volume water samples and have used the sediment retained on filter papers for analysis (e.g., Martínez-Carreras et al., 2010b; Brosinsky et al., 2014b).

Based on field observations and previous work in STCW (Koiter et al., 2013a), three main potential sources of sediment were identified for each site: (i) topsoil from agricultural fields, riparian areas adjacent to the creek and forested valley walls; (ii) unconsolidated material from surface and subsurface stream banks; and (iii) shale bedrock material exposed along the stream channel. Topsoil was collected immediately upstream from each in-stream monitoring station (Fig. 1). At each site, transects were established along several toposequences parallel to the dominant slope gradient, extending from the edge of the stream bank into the adjacent fields through the riparian area. Each transect was comprised of four to eight sampling points, representing topsoil sources from agricultural fields, riparian areas, and/or forest valley walls. At each sampling point, the surface horizon (A horizon) was sampled over its full depth (using a 3.18-cm-diameter soil probe) to characterize surface soils. In a cultivated landscape, the surface soil is regularly mixed, leading to difficulties in identifying the material “likely to be mobilized by erosion.” In fact, the likely eroded material is often applied in fingerprinting studies (e.g., Collins et al., 1997; Russell et al., 2001; Martínez-Carreras et al., 2010b). Because the identification of sampling areas that are susceptible to erosion and are well connected to the channel is subjective (Koiter et al., 2013a), a systematic approach was adopted based on the characterization of surface soils along a transect. A total of 29 composite topsoil samples were analyzed (Table 2).

Multiple stream bank samples were also collected upstream from the in-stream monitoring stations. At each site, three sampling points were defined. For each of these sampling points, stream bank samples were collected at 5- or 10-cm incremental depths along the bank profile (for a total depth ranging between 40 and 170 cm) using a box-core sampler inserted perpendicular to the stream bank face. A total of 117 stream bank samples (collected from five sites) were analyzed (Table 2).

In the Manitoba Escarpment, three representative large bedrock outcrops were sampled between sites 3 and 4 (STC-S4-OC10, 15 m high); between sites 4 and 5 (STC-S3-OC4, 8 m high), and between sites 5 and 6 (STC-S2-OC11, 13 m high) (Fig. 1). The bedrock outcrops consist of very friable shale, which enables the collection of outcrop samples similarly to the collection of stream bank samples. A total of 38 outcrop samples were collected from the different layers of bedrock (Table 2). The outcrop samples were pulverized and ground to a powder using a mortar and pestle.

Table 1. Number of suspended sediment samples analyzed for each in-stream monitoring station, 2009–2011.

<table>
<thead>
<tr>
<th>Sampling period</th>
<th>Monitoring station</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>Site 1</td>
</tr>
<tr>
<td>2009</td>
<td>May–June</td>
</tr>
<tr>
<td>July–Oct.</td>
<td>–</td>
</tr>
<tr>
<td>2010</td>
<td>Apr.–June</td>
</tr>
<tr>
<td>June–Nov.</td>
<td>1</td>
</tr>
<tr>
<td>2011</td>
<td>Apr.–June</td>
</tr>
<tr>
<td>June–Aug.</td>
<td>1</td>
</tr>
<tr>
<td>Aug.–Nov.</td>
<td>–</td>
</tr>
</tbody>
</table>
To minimize contrasts in grain size between sources and sediments, the potential source samples and the suspended sediment samples were sieved to ≤63 μm (Walling and Woodward, 1995; Walling et al., 1999; Martinez-Carreras et al., 2010b).

### Spectral Measurements and Color Coefficients Calculation

Potential source samples \((n = 184)\) and suspended sediment samples \((n = 27)\) were analyzed using a visible, near-infrared reflectance spectrometer with a spectroradiometer (ASD FieldSpec Pro, Analytical Spectral Device Inc.). Sample spectra were collected over the 350- to 2500-nm wavelength range in 1-nm increments. To perform the analyses, samples were placed on flat transparent plastic supports of ~52 mm diameter, and their surfaces were smoothed. Measurements were made on an ~3-mm layer of sample material. The samples and a white reference panel (Spectralon) were illuminated with a white light source (a halogen-based lamp [12 VDC, 20 Watt] mounted ~10 cm from the sample). Light was then collected with a fiber optic cable mounted ~2 cm from the sample/white reference panel with an angle of 45°. The angular field view of the fiber optic was 25°, resulting in an effective target area of 1.0 cm². Reflectance was computed from raw spectra obtained by the spectroradiometer using RS3 software. Absolute reflectance spectra were then derived from the CIE XYZ color space using the conversion equations defined by the CIE (Westland et al., 2012). The designations \(x\) and \(y\) specify color variations from blue to red and blue to green, respectively; \(L^*\) represents the brightness; and \(a^*\) and \(b^*\) as well as \(u^*\) and \(v^*\) represent the chromaticity coordinates as opponent red-green and blue-yellow scales (Viscarra Rossel et al., 2006). The designation CIE L’c’h* is a transformation of CIE L’\(a^*b^*\) into cylindrical coordinates. Another color coefficient, the Redness Index (RI) (Viscarra Rossel et al., 2006), was defined for the estimation of the hematite content in soils (Barron and Torrent, 1986). Because there is no single RGB color space that has achieved universal acceptance, RGB color coefficients were calculated directly from the reflectance spectra by using the conversion equations defined by the CIE. The designations \(x\) and \(y\) specify color variations from blue to red and blue to green, respectively; \(L^*\) represents the brightness; and \(a^*\) and \(b^*\) as well as \(u^*\) and \(v^*\) represent the chromaticity coordinates as opponent red-green and blue-yellow scales (Viscarra Rossel et al., 2006). The designation CIE L’c’h* is a transformation of CIE L’\(a^*b^*\) into cylindrical coordinates. Another color coefficient, the Redness Index (RI) (Viscarra Rossel et al., 2006), was computed based on the CIE L’\(a^*b^*\) color system. This index was introduced for the estimation of the hematite content in soils (Barron and Torrent, 1986). Because there is no single RGB color space that has achieved universal acceptance, RGB color coefficients were calculated directly from the reflectance spectra by averaging reflectance data in the ranges 450 to 520 nm, 520 to 600 nm, and 630 to 690 nm, corresponding to the blue, green, and red Landsat bands, respectively. The averaged reflectance data were multiplied by 255 to get the eight-bit pixel color encoding (Viscarra Rossel et al., 2006). A total of 16 color coefficients were calculated.

### Behavioral Characteristics of Color Coefficients: Prior Requirements

To be used as efficient fingerprint properties, the color coefficients should exhibit linearly additive behavior during mixing (Lees, 1997) and should exhibit conservative behavior during erosion and transport (Foster and Walling, 1994; Koiter et al., 2013a); if not conservative, then they should at least exhibit

| Table 2. Number of source samples at each sediment sampling station. Each topsoil sample is a composite of four to eight subsamples along a transect. |
|-----------------|---------------|--------------|--------------|--------------|--------------|
| Sampling site   | Field topsoil | Riparian topsoil | Forest topsoil | Stream banks | Outcrop      |
| Site 1          | 3             | 4             | –             | 12           | –            |
| Site 2          | 3             | 1             | –             | 20           | –            |
| Site 3          | 2             | –             | 2             | –            | –            |
| STC-S4-OC10     | –             | –             | –             | 13           | –            |
| Site 4          | 1             | 3             | –             | 31           | –            |
| STC-S3-OC4      | –             | –             | –             | –            | 5            |
| Site 5          | 2             | 4             | –             | 24           | –            |
| STC-S2-OC11     | –             | –             | –             | –            | 20           |
| Site 6          | 2             | 2             | –             | 30           | –            |

where \(\lambda\) is the wavelength; \(S\) is the spectral distribution of the light source; \(R\) is the spectral reflectance of the sample; and \(\mathbf{X}, \mathbf{Y}, \mathbf{Z}\) are the color-matching functions of the CIE 1964 Standard Observer, which are defined for large angular subtense (>10°) (Rochester Institute of Technology, 2013). The calculations were made with a 1-nm step increment for an A-type illuminant (a tungsten halogen–based lamp in these analyses). This lamp can be run at color temperatures in the range of ~2850 to 3300 K (Hunt and Pointer, 2011). \(K\) is a constant resulting in \(Y = 100\) for a perfect white surface.

\[
K = \frac{100}{\sum_{360 \text{nm}}^{830 \text{nm}} S(\lambda) \cdot \mathbf{Y}(\lambda)}
\]
behavior that is predictable (e.g., radioactive decay) (Belmont et al., 2014). To assess the linearly additive behavior of the 16 color coefficients, 12 artificial mixtures of source samples were prepared by mixing two to four sources chosen randomly among the potential source samples. Similarly to source samples, mixture samples were analyzed by visible, near-infrared reflectance spectrometry, and the 16 color coefficients were computed from the absolute reflectance data. Mixture sample color coefficients were then compared with the color coefficients of the source samples used to prepare the mixtures. By computing the RMSE using the predicted color coefficient (linear combination of the color coefficients of the sources used for the mixture weighted by the proportions), the measured color coefficient (color coefficient of the mixture sample), and the number of samples, it was possible to assess the linearly additive behavior of each color coefficient. Normalized RMSE values by the mean of each color coefficient are reported to eliminate the scale difference among color coefficients.

The conservative behavior of color coefficients was not tested in the present study. A recent study (Legout et al., 2013) has shown that colorimetric parameters L*, a*, b*, u*, and v* generally exhibit low variation according to the immersion time of source samples (black marls, molasses, limestone) in river water (0–70 d). However, these authors also observed that black marls and molasses may be more sensitive than limestone to colorimetric variations. That research highlights the need for further investigation about the possible transformations of color coefficients during transport and erosion (e.g., oxidation of iron and decomposition of organic matter) (Koiter et al., 2013b; Brosinsky et al., 2014a).

### Source Classification

Based on Koiter et al. (2013a), we collected samples to represent three broad source types: (i) topsoil (field, riparian, and forest valley walls), (ii) channel banks, and (iii) shale outcrops. Subsequently, these three types were divided into nine source groups in an attempt to reflect the physiography of the watershed and to provide the greatest amount of information on the location of the dominant sources (Table 3). The discriminatory power of stable isotope mixing models decreases with an increasing number of sources, particularly when the number of sources is greater than six (Phillips et al., 2014). To keep the number of sources as low as possible without excluding potential sources, we performed canonical discriminant analysis using Proc Candisc of SAS (SAS Institute Inc., 2008) to reclassify the nine original source groups (Table 3) into a smaller number of sources (i.e., fewer than six) (Phillips et al., 2014).

### Color Coefficients Selection for Source Discrimination

A nonparametric Kruskal–Wallis H-test was performed using open source statistical software R packages for each color coefficient to evaluate the existence of intercategory contrasts among the source samples (Collins and Walling, 2002). The color coefficients showing no intercategory contrasts among the source samples ($P > 0.05$) were eliminated from the dataset. Because the Stable Isotope Analysis in R (SIAR; v 4.2) model (Parnell et al., 2010; discussed in detail below) assumes that each fingerprint property of both source and sediment is “reasonably” normally distributed, each individual parameter was tested for normal distribution for each source using the Shapiro–Wilk test. Violating the assumption of normality may result in a misrepresentation of the fingerprint variability in the SIAR mixing model; thus, color variables violating normality were excluded in the further analyses. A discriminant function analysis (DFA) and a stepwise DFA were performed on the normally distributed color variables to select composite fingerprints for the SIAR model. A DFA was used to assess the individual discrimination power of each or multiple color coefficient by determining the number of samples correctly classified into the source groups using leave-one-out cross validation (Reimann et al., 2008). After investigating the individual discrimination power of each color coefficient, a forward stepwise DFA based on the Wilk’s lambda criterion was performed to determine the best combination of variables (color coefficients) in discriminating source groups without redundant information. Discriminant function analysis and stepwise DFA (niveau $= 0.05$) were performed using R packages (MASS and klaR) (Venables and Ripley, 2002; Weihs et al., 2005).

### Source Ascription Modeling

The source contributions to the sediment load were estimated at each in-stream monitoring station using SIAR (v 4.2) available as an open source R package developed for ecological studies (Parnell et al., 2010). Recently, SIAR has been used to determine the sources of sediment using geochemical and radionuclide fingerprints (e.g., Dutton et al., 2013; Koiter et al., 2013a). The model assumes that each sediment fingerprint property comes from a Gaussian distribution with an unknown mean and standard deviation. The weights are source proportions, which are given a Dirichlet prior distribution. The Dirichlet prior...
distribution default given by SIAR was used; it is designed to be vague so that the results are primarily influenced by the data. Model fitting is via Markov chain Monte-Carlo, which produces simulations of plausible values of the relative contributions of the different sources.

Because stream bank materials within and below the escarpment were substantially different than above the escarpment, they were treated as two separate sources. In practice, sediment sources in the lower reaches should not contribute sediments sampled in the upper reaches. Therefore, stream bank materials within and below the escarpment (source 3) were excluded in the model when determining the source contribution at sites 1 and 2, which are located above the escarpment. The source contribution to the suspended sediment at each station was estimated separately using the SIAR model. The model was run using 200,000 iterations with a burn-in of 50,000. The trophic enrichment and concentration dependence factors, used for ecological study Parnell et al., 2010), were omitted, in agreement with other similar studies (Koiter et al., 2013a). Some studies have included particle size, organic matter, and conservativeness correction factors in the mixing model formulation (e.g., Walling and Woodward, 1995; Collins et al., 1997; Walling et al., 1999; Gruszowski et al., 2003; Motha et al., 2003). However, regarding the complexity of the relationships between tracer concentration and grain size (Russell et al., 2001) as well as organic matter content highlighted by these studies, no correction factors were applied, following the recommendations of Martinez-Carreras et al. (2010a), Smith and Blake (2014), and Verheyen et al. (2014).

**Results and Discussion**

**Linearly Additive Behavior of Color Coefficients**

The normalized RMSE values by the mean for each color coefficient are presented in Fig. 4. The x, y, and h coefficients showed strong linearly additive properties with the lowest normalized RMSE (<0.05). Most of the color coefficients showed linearly additive properties with the normalized RMSE <0.2. However, RI is the only color coefficient that exhibited exceptionally high normalized RMSE, suggesting a nonlinear effect. This nonlinear effect increases with the number of samples used to prepare the mixture (data not shown) and can be due to undetermined interaction effects (i.e., changing physical characteristics of the source materials once mixed together). This nonadditivity phenomenon was observed with mineral magnetic properties, especially when sources are strongly ferromagnetic (Lees, 1997). Therefore, RI cannot be used as a fingerprint property and was removed from the dataset before analysis of source discrimination and source ascription modeling. These findings suggest that sediment sources can be “unmixed” using linear modeling based on the color parameters x, y, X, Z, Y, L, a, b, u, v, c, h, R, G, and B.

**Source Type Discrimination**

The results of the canonical discriminant analysis on the source samples are shown in Fig. 5. The first two discriminant functions explain 93% of the overall variability before source aggregation and explain 97% after source aggregation, suggesting that most of the information on the source samples is reflected by plotting the first discriminant function against the second discriminant function. On this two-dimensional scatterplot (Fig. 5b), the outcrop shale materials (A1) are easily distinguishable from the other three sources. Both the outcrop shale materials and stream bank materials within and below the escarpment originated from the local sources, explaining the close relationship between these two sources, as shown in Fig. 5. The topsoil above the escarpment is different from the topsoil within and below the escarpment because the former is partially from external sources as a result of ancient glacial deposits. However, the topsoil cannot be distinguished based on physiography and origins (e.g., above and below the escarpment) as indicated by the scatter plot of the first two discriminant functions, suggesting
that physiography and origins were not the key factors in separating topsoil in this study. Soil organic matter content is closely linked to soil color (Pulleyman et al., 2000). Agricultural practices might dramatically change soil organic matter content, making it difficult to discern differences in this property across the watershed. Organic matter content might override the physiography and origin, resulting in the lack of topsoil separation among topsoil subsources. The erosion of hilltops in the watersheds leaves loose shale materials exposed and affects the soil color, explaining high variation and scattered distribution among different sources of topsoil. According to the topsoil distribution pattern in Fig. 5, it is reasonable to combine topsoils from the riparian area (B4), the forest valley walls (part of C4), and the agricultural fields (B3 and C3) into one source.

The results of the canonical discriminant analysis highlighted a lack of dimensionality in the dataset. Walden et al. (1997) suggested a small number of source types for efficient source type ascription. Similarly, reclassifying and downsizing original source groupings has been used in other color fingerprinting studies (e.g., Brosinsky et al., 2014a, 2014b; Verheyen et al., 2014). In the current study, where nine original sources were originally identified, it is necessary to reduce the number of source types by combining the sources with similar color fingerprints. According to the patterns shown in Fig. 5, four source groups were considered for further analysis: group 1, topsoil from field, riparian area, and forest valley walls (B3, B4, C3 and C4); group 2, stream bank materials above the escarpment (C1 and C2); group 3, stream bank materials within and below the escarpment (B1 and B2); and group 4, outcrop shale materials (A1).

The results of the Kruskal–Wallis H-test (Table 4) indicated that each color coefficient was different among the four source types ($P < 0.01$), suggesting that the hypothesis of stochastic homogeneity is not true for at least one source (Ruxton and Beauchamp, 2008). However, those results did not confirm significant differences between all pairs of source category (Collins and Walling, 2002), and further assessment of the source discrimination is required. The results of the DFA using all 10 variables ($X, Z, Y, L, B, a, u, R, G$, and $x$) that pass the normality test showed that 64 to 74% of the samples was classified into the correct source category (Table 4). The stepwise discriminant analysis indicated that four variables of $Z, L, X, Y$ provided an optimum combination in discriminating sources, with 83% of overall samples being correctly classified. The remaining six color variables of $B, a, u, R, G$, and $x$ were excluded in the SIAR model because including any of them reduced the percentage of samples being correctly classified (data not shown). The results of the statistical analysis clearly demonstrated that combining certain color properties improved source discrimination. Using the four selected color coefficients, the correctly classified percentage was 66% for topsoil samples, 67% for stream bank materials collected above the escarpment, 89% for stream bank materials within and below the escarpment, and 97% for outcrop shale materials (Table 5). These results showed that source 3 (stream bank materials within and below the escarpment) and source 4 (outcrop shale materials) were sufficiently well defined and that the color parameters provided acceptable source discrimination for source ascription modeling within STCW. However, as expected because of the high variability of the field, riparian, and forest topsoils and the similarity among these topsoils, the topsoil source was poorly defined, with approximately 66% of source samples being correctly classified. Increasing the dimensionality of the dataset by considering other fingerprint properties (e.g.,

Table 4. Results of the Kruskal–Wallis H-test and the discriminant function analysis applied to the measurements of selected color-based fingerprints made on the potential source materials.

<table>
<thead>
<tr>
<th>Fingerprint property</th>
<th>Kruskal–Wallis p value</th>
<th>Percentage of samples correctly classified (discriminant function analysis)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>X</td>
<td>&lt;0.01</td>
<td>74</td>
</tr>
<tr>
<td>Z</td>
<td>&lt;0.01</td>
<td>68</td>
</tr>
<tr>
<td>Y</td>
<td>&lt;0.01</td>
<td>71</td>
</tr>
<tr>
<td>L</td>
<td>&lt;0.01</td>
<td>72</td>
</tr>
<tr>
<td>B</td>
<td>&lt;0.01</td>
<td>69</td>
</tr>
<tr>
<td>a</td>
<td>&lt;0.01</td>
<td>69</td>
</tr>
<tr>
<td>u</td>
<td>&lt;0.01</td>
<td>64</td>
</tr>
<tr>
<td>R</td>
<td>&lt;0.01</td>
<td>71</td>
</tr>
<tr>
<td>G</td>
<td>&lt;0.01</td>
<td>70</td>
</tr>
<tr>
<td>x</td>
<td>&lt;0.01</td>
<td>65</td>
</tr>
</tbody>
</table>

† Source 1: topsoil form agricultural field, riparian areas, and forest valley walls (B3-B4-C3-C4); source 2: stream bank materials above the escarpment (C1-C2); source 3: stream bank materials within and below the escarpment (B1-B2); and source 4: outcrop shale materials (A1).

Table 5. Results of the forward stepwise discriminant function analysis based on the Wilk’s lambda criterion to determine the combination of fingerprint properties allowing most of discrimination between sources.

<table>
<thead>
<tr>
<th>Fingerprint property</th>
<th>Wilk’s lambda</th>
<th>Partial p value</th>
<th>Percentage of samples correctly classified†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Z</td>
<td>0.347</td>
<td>&lt;0.01</td>
<td>Z</td>
</tr>
<tr>
<td>L</td>
<td>0.213</td>
<td>&lt;0.01</td>
<td>Z + L</td>
</tr>
<tr>
<td>X</td>
<td>0.191</td>
<td>&lt;0.01</td>
<td>Z + L + X</td>
</tr>
<tr>
<td>Y</td>
<td>0.174</td>
<td>&lt;0.01</td>
<td>Z + L + X + Y</td>
</tr>
</tbody>
</table>

† Source 1: topsoil form agricultural field, riparian areas, and forest valley walls; source 2: stream bank materials above the escarpment; source 3: stream bank materials within and below the escarpment; and source 4: outcrop shale materials.
geochemistry, radionuclides, and/or organic compounds) could improve the source discrimination (Collins and Walling, 2002; Martinez-Carreras et al., 2010a), suggesting that integration of color fingerprint properties and conventional fingerprint properties might provide an ideal approach to discriminate topsoil subsources in the study watershed.

**Source Type Ascription**

To assess the potential of the color parameters to identify the dominant source of the suspended sediment, the distribution of the suspended sediment samples collected along the main stem of the South Tobacco Creek as well as the means and standard deviations of the potential sources are shown in Fig. 6. Source 4, the outcrop shale materials, was clearly distinguished from other three sources; in contrast, topsoil (source 1) and stream bank materials above the escarpment (source 2) are closely associated. This is consistent with the results of the stepwise discriminant analysis, with a high percentage of source 4 being correctly classified and low percentages of sources 1 and 2 being correctly classified. All suspended sediment samples lie within the range of the potential source materials (Fig. 6), suggesting that all major potential sources were included in the model and that the alteration of the composition of suspended sediments during transport and storage (i.e., changes in particle size composition) was relatively small (Walden et al., 1997).

The SIAR mixing model was applied to the four variables chosen by the stepwise DFA to establish the relative contribution of each source to suspended sediment collected from 2009 to 2011. The results of the source ascription modeling, including uncertainty assessment, are presented in Fig. 7. The proportions were reported as the credibility intervals ranging between the 25th and 75th percentiles. The SIAR model outputs confirmed the switch in the predominant sources between the headwaters and the watershed outlet, which was highlighted by Koiter et al. (2013a). In the low-order streams in the headwaters at sites 1 and 2, topsoil are the predominant

![Fig. 6. Biplot of selected color variables (L, X, Y, and Z) for the suspended sediment samples and the potential source samples. Source 1: topsoil from agricultural fields, riparian areas, and forest valley walls; source 2: stream bank materials above the escarpment; source 3: stream bank materials within and below the escarpment; and source 4: outcrop shale materials.](image-url)
source of sediment, accounting for approximately 41 to 60% and 38 to 60%, respectively. The model identified that shale bedrock materials also contributed to the sediment load in the headwaters, accounting for approximately 12 to 19% (site 1) and 8 to 18% (site 2). Although we did not collect bedrock shale samples above the escarpment, these results are in accordance with the field observations because some small outcrops were observed above the escarpment; furthermore, there is no reason to expect the color material properties for these unsampled small outcrops to differ from the large ones sampled (i.e., A1). In many respects, these results suggested that the unmixing model approach provided a realistic determination of the sediment sources based on our understanding of the sediment dynamics in the upper watershed.

As the stream flows through the escarpment, the contribution of the topsoil decreased (19–41% for site 3 and <3% for sites 4, 5, and 6), whereas the contribution of the bedrock outcrop materials increased, with maximum values of 68 to 77% for site 4. At the bottom of the escarpment, the contribution of the bedrock materials was reduced (53–62% for site 5 and 60–69% for site 6), but the contribution of the stream bank materials increased, accounting for approximately 8 to 22% (site 4), 17 to 35% (site 5), and 10 to 26% (site 6). These results confirmed the switch to a dominance of stream banks and shale bedrock sources in the middle and lower reaches, as reported by Koiter et al. (2013a), which is in keeping with the obvious signs of bank erosion in the downstream reaches. However, these results are slightly different from the findings by Koiter et al. (2013a). In particular, the use of the color-based fingerprinting approach showed a smaller contribution of topsoil sources. This may reflect the fact that the methodology used in the two studies differs. First, the number of samples collected in the current study, particularly for the outcrop samples, was increased, allowing a better description of the outcrop source. Second, source classification was also defined differently to take into account source heterogeneity as the stream flows through the escarpment, leading to changes in the proportion results. Third, the color-based fingerprinting approach was relatively weak in discriminating between the topsoil and stream bank materials above the escarpment due to the effects of glacial deposition. For both approaches, the results involve high uncertainties, which were reduced considerably by the statistical technique used and better source classification, reflecting source heterogeneity throughout the watershed. Most importantly, both approaches provide similar conclusions and are capable of demonstrating the switch in the sources of sediments found within the watershed, which is consistent with our understanding of the geomorphology of the watershed. To improve the prediction of each source contribution and to reduce uncertainties, we recommend combining the two techniques in a composite fingerprinting framework (e.g., Collins et al., 1997; Walling et al., 1999).

The role of the Manitoba Escarpment on sediment sources is highlighted by the increasing contribution from shale bedrock materials to the suspended sediment load as the creek incises into this material. The small contribution of topsoil sources may be explained by the construction of small dams and reservoirs in the headwaters, which have been shown to be net sediment sinks (Tiessen et al., 2011) because the topsoil-rich sediment may be stored in these water retention structures. Floodplains and beaver dams may also play a role in this storage of topsoil-rich sediment. These findings have important implications in terms of which BMPs may be most suitable for reducing the delivery of sediment and associated P to downstream waterbodies, such as Lake Winnipeg. The findings suggested that BMPs that reduce soil erosion and sediment delivery from fields to channels may be more suited to upper parts of watersheds, whereas BMPs that reduce bank erosion may be more suited to lower parts of watersheds.

Conclusions

Color fingerprints offer substantial information on sediment dynamics at the watershed scale. As with the conventional fingerprinting approach from a previous study, the color-fingerprinting approach highlighted a switch in the predominant sources from the headwaters to the watershed outlet. The findings showed that topsoil was the main sediment source in the upper parts of the study watershed, whereas outcrop shale materials and stream banks were the main sources of sediment in the lower parts of the watershed. This research shows the importance of testing the behavior of color-based tracers and appropriate grouping of sediment sources when assessing the sources of sediment. The color fingerprinting approach can be integrated in a composite fingerprinting framework, combining color fingerprints and conventional fingerprints (e.g., geochemical and fallout radionuclide properties), to improve source discrimination, which is a key component for source ascription modeling. The use of color fingerprints is seen as a promising, cost-effective technique for sediment source fingerprinting.

Acknowledgments

This work is funded by a NSERC (Natural Science and Engineering Research Council of Canada) Strategic Grant—Development of environmental fingerprinting techniques for sources of sediment and associated phosphorus within agricultural watersheds of Canada.

Fig. 7. Box and whisker plots showing the relative contributions of potential sources to suspended sediment collected from the monitoring stations. Source 1: topsoil from agricultural fields, riparian areas, and forest valley walls (B3-B4-C3-C4); source 2: stream bank materials above the escarpment (C1-C2); source 3: stream bank materials within and below the escarpment (B1-B2); and source 4: outcrop shale materials (A1). The lower and upper hinges correspond to the first and the third quartiles, respectively. The whiskers are created using the Tukey method. Potential outliers are plotted as crosses.


