

## Approach to assess agroecosystem anthropic disturbance: Statistical monitoring based on earthworm populations and edaphic properties



C. Masin<sup>a</sup>, A.R. Rodríguez<sup>b</sup>, C. Zalazar<sup>a</sup>, J.L. Godoy<sup>a,c,\*</sup>

<sup>a</sup> INTEC (CONICET and UNL), Güemes 3450 (3000), Santa Fe, Argentina

<sup>b</sup> Universidad Nacional del Litoral – Facultad de Humanidades y Ciencias, Santa Fe, Argentina

<sup>c</sup> FRP-UTN (Facultad Regional Paraná – Universidad Tecnológica Nacional), Paraná, Argentina

### ARTICLE INFO

#### Keywords:

Agricultural systems  
Land use and management practices  
Anthropic disturbance level  
Earthworms  
Bioindicators  
Edaphic properties

### ABSTRACT

Land degradation due to anthropic factors is the reduction of its actual or potential productivity. Nowadays, this topic is a major concern, as it affects more than one third of the soil in the world. This work presents an empirical assessment of the anthropic disturbance level (ADL) for agricultural and livestock production systems. This assessment is obtained by mapping the characteristics of land use and management practices by using five specific indicators and integrating them into a global indicator (ADL score). Earthworm populations (good indicators of soil quality) in soils under different production systems are studied to determine if the population changes are attributable to the intensity of land use and management practices. A correlation model between ADL, edaphic properties, and earthworm population characteristics is developed by using samples of 20 sites in Santa Fe province, Argentina. The inclusion of ADL allowed finding a consistent correlation structure. The results also showed that earthworm density, species diversity, and activity change at the different sites were highly sensitive to anthropic disturbance. Based on this data-driven model, the ADL can be estimated by measuring edaphic and biological data on a soil sample to monitor soil conditions for different production systems. Thus, ADL monitoring would allow deciding how to continue using and managing the land to improve its sustainability.

### 1. Introduction

Over the last four decades, agricultural-livestock production in Argentina has expanded noticeably, covering large regions of the country (Carreño and Viglizzo, 2010). In this sense, Santa Fe province underwent major transformations in the landscape heterogeneity of its phytogeographical regions (Miretti et al., 2012). For agroecosystems, the intensification of production practices (tillage, agrochemicals, cattle grazing, crop rotation, replacement of natural vegetation, irrigation system, etc.) leads to the modification of structural, nutritional and biological characteristics of the soil (Ernst and Emmerling, 2009), thus reducing ecosystem services (Johnston et al., 2018).

In recent decades, several biological indicators have been proposed to evaluate the sustainability of soil use, for which edaphic macroinvertebrates were studied (Parisi et al., 2005; Culman et al., 2010; Dewi and Senge, 2015; Rudisser et al., 2015). Among the bioindicators used to evaluate the sustainability of soil use in agroecosystems, earthworms are one of the most frequently studied (Lavelle et al., 2007; Pérès et al., 2011; Paoletti et al., 2013; Bartz et al., 2014; Carnovale

et al., 2015; Falco et al., 2015; Roarty et al., 2017).

Earthworms are ecosystem engineers, since they can modify soil structure and fertility directly through digestion and burrowing activities, and indirectly by encouraging other beneficial soil organisms (Jones et al., 1994; Blouin et al. 2013). The overall effect of earthworms on soil processes depends on the ecological category of the particular species. The three main categories are: epigeic, anecic and endogeic (Bouché, 1977), although intermediate categories can be used (e.g. epiendogeic and endoanecic) (Cunha et al., 2016). Epigeics are litter transformers, which mineralize and digest plant remains at the soil surface; anecics feed on plant material usually mixed with soil, and generate burrows in a more vertical direction; endogeics are geophagous species that form temporary horizontal burrows in the mineral soil and feed on soil organic matter of different qualities (Lavelle, 1981). The soils used for agricultural activities over a long period show edaphic conditions to which certain earthworm species are usually tolerant and, therefore, they become dominant, such as some endogeics of the genus *Aporrectodea* (Kherbouche et al., 2012).

Studies developed in Argentina (Masin et al., 2011; Maitre et al.,

\* Corresponding author at: INTEC (CONICET and UNL), Güemes 3450 (3000), Santa Fe, Argentina.

E-mail address: [jlgodoy@santafe-conicet.gov.ar](mailto:jlgodoy@santafe-conicet.gov.ar) (J.L. Godoy).

2012; Bedano and Domínguez, 2016; Domínguez et al., 2018) and in other regions of the world (Bartz et al., 2013; Lemtiri et al., 2014; Bertrand et al., 2015; Kanianska et al., 2016; Xie et al., 2018) show that earthworm populations are conditioned not only by the edaphic properties and the type of vegetation but also by the history and intensity of land use. In this sense, the studies by Butt and Lowe (2004), García-Pérez et al. (2014), and Ortiz-Gamino et al. (2016) report that, while earthworm populations are conditioned by some edaphic properties, they are highly sensitive to the anthropic disturbance generated in their ecosystems, thus providing valuable information as bioindicators of soil quality (or soil health). In particular, Bertrand et al. (2015) reported that the composition of earthworm populations (number of individuals, proportion of juveniles or adults) depends on both production system and tillage intensity and frequency, and also that population density is affected by changes in vegetation, crop residues quantity and quality, and grazing activities.

The edaphic properties most related to earthworms are: soil organic matter (SOM) content, bulk density (BD), cation exchange capacity (CEC), soil texture, and pH. For example, a BD lower than the critical value (1.3 [g/cm<sup>3</sup>]) does not imply that soil quality is optimum for the development of earthworms, because this is also conditioned by the type of tillage, frequency of application of agrochemicals, livestock overstocking, etc. (Álvarez et al., 2012). In soils with more than 10 years of agricultural use, a BD between 1.21 and 1.26 [g/cm<sup>3</sup>] was recorded and earthworm abundance was not affected in any case, but with a BD from 1.38 [g/cm<sup>3</sup>], the abundance decreased abruptly (Domínguez et al., 2009). The relationship between earthworms and some chemical properties (pH and available nutrients) was also studied by Kanianska et al. (2016), as well as the relationship between earthworm density and diversity and land use, management practices, and selected abiotic and biotic indicators.

Thus, tools are needed to better predict how land management practices affect the provision of ecosystem services through their effects on soil biota. In this sense, Johnston et al. (2015) presented a mechanistic model (based on literature data) for predicting how agricultural management practices (pesticide applications and tillage) affect soil functioning (crop yields) by affecting earthworm populations (earthworm biomass). Fusaro et al. (2018) proposed an index called QBS-e (Soil Biological Quality Index), which is similar to previous ones by Parisi (2001) and Paoletti et al. (2013). The QBS-e index is based on earthworm ecological categories, not on species. This index enables the estimation of the sustainability of soil management practices, and it can be used also by non-experts and directly in the field. However, QBS-e index can have some limitations as edaphic properties are not taken into consideration in this index.

In spite of the studies already conducted, it is still necessary to establish quantitative relationships between these three factors (biological, edaphic, and agricultural management data). In particular, it is important to consider the anthropic disturbance level of an agroecosystem in order to find a valid correlation structure between edaphic properties and earthworm populations.

The assessment of the health of agroecosystems still requires validated tools, able to inform state agencies about the degradation and/or remediation of soil properties and functions. It is important that control agencies have validated tools for supervising land use and management practices in order to maintain soil sustainability. To our knowledge, no method for assessing the anthropic disturbance levels in an agroecosystem has been reported yet. In particular, there are no validated tools to identify and prevent the degradation of an agroecosystem, based on the relationships between earthworm populations, edaphic conditions, and agricultural management practices.

In this work, an empirical assessment of the anthropic disturbance level (ADL) for agricultural and livestock production systems is presented. It is based on scoring land use and management practices. Then, an ADL estimator is developed based on a multivariate correlation model, which was calibrated with data from 20 sampled agricultural

sites from Santa Fe province (Argentina). This estimator is an inferential model that uses the available measurements to estimate (or predict) the ADL value of a site. In particular, the ADL estimator takes into account edaphic and biological data and is a valuable methodology to monitor the health of agroecosystems. This methodology can be a promising tool to decide how to continue using and managing the land to improve its sustainability. Also, the model can be calibrated with data from any other region of interest (different from those used here for calibration), which would allow the application to be used for any region of the world.

## 2. Materials and methods

### 2.1. Empirical assessment of the ADL in agricultural and livestock production systems

In order to quantify the impact of certain agricultural-livestock activities, this work proposes to characterize land use and management practice types through specific indicators to subsequently integrate these indicators to assess the overall level of anthropic disturbance. These specific indicators also allow understanding the sustainability state of a production system and its risk factors (Sarandón, 1998). These proposed indicators assess the following characteristics: *I*<sub>1</sub>- land cover (natural or cultivated); *I*<sub>2</sub>- land use type (or productive activity type); *I*<sub>3</sub>- soil tillage history (tillage type and time); *I*<sub>4</sub>- crop diversity and rotation over time; *I*<sub>5</sub>- fertilizers and pesticides (number of annual applications and doses). Each specific indicator (*I*<sub>1</sub>, *I*<sub>2</sub>, *I*<sub>3</sub>, *I*<sub>4</sub>, *I*<sub>5</sub>) is subdivided into several classifications and their respective scores to evaluate the effect (or sustainability) that this type of management practice exerts on the production system (or agroecosystem). Table 1 shows -for each indicator- the proposed mapping between classifications and their assigned empirical disturbance scores. Once each indicator is assessed, the ADL score (or the overall indicator score) is obtained by averaging the specific indicators as follows:

$$ADL = (I_1 + I_2 + I_3 + I_4 + I_5)/5 \quad (1)$$

### 2.2. Data collection

The samples were collected in different agricultural and livestock production systems (or sites) to study the effect of land use and management practices on edaphic properties and earthworm populations. For this purpose, edaphic properties were measured, earthworm populations were counted, and the backgrounds of land use and management practices were assessed for each site.

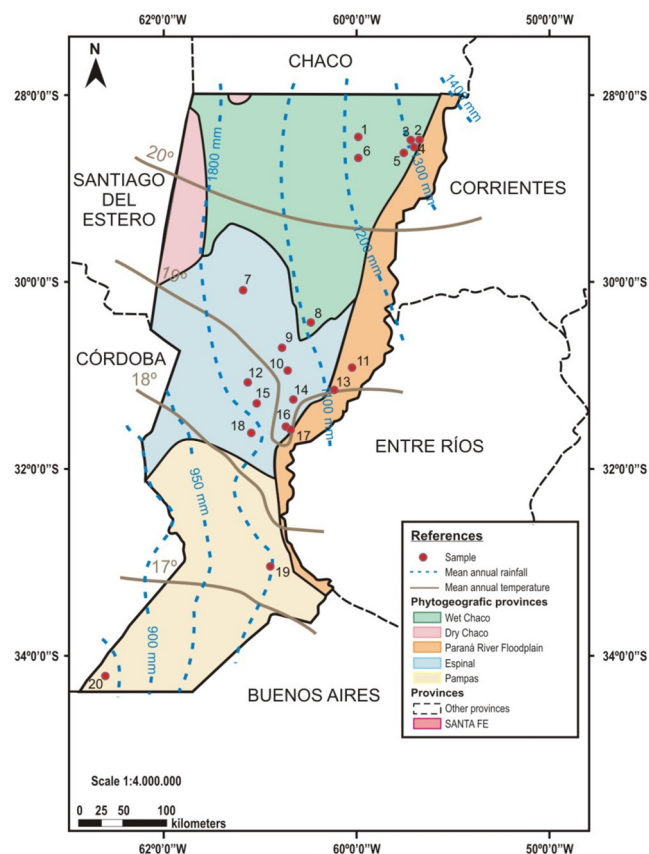
#### 2.2.1. Geographic area of the study

The study was conducted in Santa Fe province (see Fig. 1), Argentine Republic, which is located between 59° and 63° W and between 28° and 34° S. This province -with a total area of 133007 km<sup>2</sup>, soft relief and differentiated climates- is an extensive plain that ranges from 10 to 125 m above sea level (Biasatti et al., 2016). The climate of Santa Fe has two gradients, one thermal, from north to south, and another hydrological, from east to west (varying from humid to subhumid). Thus, the north of the province has a warm climate, with average temperatures of 21 °C and with annual precipitations ranging from 950 mm in the west to more than 1300 mm in the east (Lewis and Collantes, 1974). In the south of the province, mild climate with Pampas characteristics prevails (it does not register extreme heat nor intense cold) and humidity is high due to rainfall, which is more abundant in summer. The main types of vegetation in Santa Fe are comprised in four phytogeographical regions (Burkart et al., 1999): 1) Chaco region, with two sub-regions: 1.A) Dry Chaco region: located in the northwest and characterized by water deficit with predominance of xerophile forests. 1.B) Wet Chaco region: located in the northeast and north-central, with more than 1000 mm annual rainfall. Its vegetation includes humid subtropical deciduous

**Table 1**  
Subdivision (with empirical scoring) of the specific indicators of the overall assessment of ADL in Santa Fe agroecosystems.

Specific Indicators	Anthropic Disturbance Gradient ( - → + )							
	0	Low ADL Score interval: [0 – 2.5)			Medium ADL Score interval: [2.5 – 3.5)			High ADL Score interval: [3.5 – 5)
<i>I</i> <sub>1</sub> : Ground cover (Score)	Untouched natural vegetation (0)	Natural vegetation minimally intervened (1)			Implanted vegetation and crop stubble (3)			Minimum coverage by crop stubble (4)
<i>I</i> <sub>2</sub> : Land use type (Score)	None (0)	Extensive livestock <sup>1)</sup> (2)			Family horticulture (2.5)	Agriculture / Livestock (3)	Intensive horticulture (3.5)	Intensive agriculture (4.5)
<i>I</i> <sub>3</sub> : Tillage type and time (Score)	None (0)	NT <sup>2)</sup> (≤5 years) (0.5)	NT (≤10 years) (1)	NT (≤15 years) (2)	NT (≤20 years) + MT <sup>3)</sup> (≤10 years) (3)			NT (>25 years) + MT (≤20 years) (4.5)
<i>I</i> <sub>4</sub> : Diversity of crops and rotations per year (Score)	Natural vegetation (0)	Polyculture with rotations (1)		1 crop with soil rest (1.5)	3 crops with rotation (2)	2 crops with rotation (3)		1 crop without rotation (3.5)
<i>I</i> <sub>5</sub> : Use of fertilizers and/or pesticides (Score)	Not used (0)	Organic fertilizer by livestock grazing (0.5)	Organic fertilizer (1)	Organic fertilizer + chemical fertilizer w/RD <sup>4)</sup> (1.5)	Pesticides w/LD <sup>5)</sup> (2)	Organic fertilizer + fertilizers w/RD + pesticides w/LD (2.5)	Fertilizers w/RD + pesticides w/HD <sup>6)</sup> (3)	Fertilizers w/HD + pesticides w/HD (4.5)
<b>Overall Indicator:</b> $ADL = (I_1 + I_2 + I_3 + I_4 + I_5) / 5$								

- 1) This value is weighted by 5 when there is livestock overstocking, i.e.  $I_2 = 2*5$ .
- 2) NT: Non Tillage (or No-till direct seeding).
- 3) MT: Minimum Tillage.
- 4) RD: Recommended Dose.
- 5) LD: Low Dose.
- 6) HD: High Dose.



**Fig. 1.** Location of the 20 sampled sites latitudinally numbered. Phytogeographical regions, isotherms, and isohyets.

forests, palms and grasslands with wetlands. 2) Espinal region: located in the middle and characterized by the presence of low xerophile forests. 3) Paraná river floodplain region: with subtropical wet forest, gallery forest, savannas and wetlands (rivers, streams, ponds, marshes and estuaries). 4) Pampas region: in the south, it is mainly composed by different types of grasslands.

The 20 sites sampled throughout the province (Fig. 1) have different histories of land use and agricultural management practices. The agricultural use of the land is classified into: Livestock in Woodland (LW), Livestock in Grassland (LG), Agriculture (A), Agriculture/Livestock (A/L), or Horticulture (H). Based on this classification, the sites analyzed are renamed according to the following nomenclature: use classification and latitudinal order (Table 2).

**2.2.2. Earthworm and soil sampling**

The earthworms were sampled through hand sorting from standard soil volumes (Jiménez et al., 2006). Sampling was conducted in two seasons (autumn and spring) between 2015 and 2017 in 11 out of the 19 districts of province. In order to have a representative estimation of the abundance of each species in each site, 20 soil monoliths of 30 × 30 × 30 [cm], spaced 15 [m] apart, were collected in two seasons (i.e., 40 soil monoliths per site). The taxonomic identification of each specimen at species level was performed according to Mischis (1991) and Reynolds (1996). The different species were grouped into three ecological categories: Epigeic (1), Epiendogeic (2), and Endogeic (3), as they respond differently to the land use and management practice types. Ecological categories provide information about the function that an earthworm performs in its habitat, based on eating habits, behavior, morphology and demography. In this work, the earthworm populations of each site are characterized by the following variables: population total density ( $D_T$  [individuals/m<sup>2</sup>]), percentage of adults (%Ad), species diversity ( $H$ ) -estimated by Shannon diversity index- (Southwood and Henderson, 2000) and density for each category: Epigeic ( $D_1$ ), Epiendogeic ( $D_2$ ) and Endogeic ( $D_3$ ), where  $D_T = D_1 + D_2 + D_3$ .

The soil of each site was sampled by mixing three soil monoliths to determine its edaphic properties: content of SOM, by Walkley and Black

**Table 2**  
Surveys carried out to farmers to characterize and assess the anthropic disturbance of each site sampled in Santa Fe province.

Latitudinal order	Site-Land use type	Geographic localization	Condition surveyed/registered in each specific indicator	ADL score (dimensionless)
1	LW <sub>1</sub>	28°27'07.13" S; 59°58'57.59" W	I <sub>1</sub> : NVP <sup>1</sup> with minimum extraction of trees. I <sub>2</sub> : ELNW <sup>2</sup> for 25 years. I <sub>3</sub> : None. I <sub>4</sub> : Natural vegetation. I <sub>5</sub> : No use of agrochemicals, only natural organic fertilizer (cattle dung) by livestock grazing.	ADL = (1 + 2 + 0 + 0 + 0.5)/5 = 0.7 (Low ADL)
2	A <sub>1</sub>	28°25'22.64" S; 59°22'39.17" W	I <sub>1</sub> : Minimum ground cover by crop stubble. I <sub>2</sub> : Intensive agriculture for 50 years. I <sub>3</sub> : NT (greater than 25 years) plus MT (20 years). I <sub>4</sub> : 3 crops w/R <sup>3</sup> . I <sub>5</sub> : Chemical fertilizers w/RD and pesticides w/HD.	ADL = (4 + 4.5 + 4.5 + 2 + 3)/5 = 3.6 (High ADL)
3	A/L <sub>1</sub>	28°27'33.11" S; 59°18'58.00" W	I <sub>1</sub> : Ground cover by implanted vegetation (pastures) and crop stubble. I <sub>2</sub> : Agriculture and livestock for 25 years. I <sub>3</sub> : NT (last 20 years) plus MT (last 10 years). I <sub>4</sub> : 3 crops w/R. I <sub>5</sub> : Chemical fertilizers w/HD and pesticides w/HD.	ADL = (3 + 3 + 3 + 2 + 4.5)/5 = 3.1 (Medium ADL)
4	H <sub>1</sub>	28°29'16.28" S; 59°20'55.40" W	I <sub>1</sub> : Year-round vegetation cover with season crops. I <sub>2</sub> : Family horticulture for 28 years. I <sub>3</sub> : NT in last 25 years plus MT (20 years) and. I <sub>4</sub> : Polyculture (different crops) w/R. I <sub>5</sub> : No use of agrochemicals, only natural organic fertilizer (cow and chicken manure). Biological control of pests.	ADL = (3 + 2.5 + 4.5 + 1 + 1)/5 = 2.4 (Low ADL)
5	A <sub>2</sub>	28°37'19.13" S; 59°30'43.20" W	I <sub>1</sub> : Minimum ground cover by crop stubble. I <sub>2</sub> : Intensive agriculture for more than 30 years. I <sub>3</sub> : NT in the past 25 years plus MT (20 years). I <sub>4</sub> : 3 crops w/R. I <sub>5</sub> : Chemical fertilizers w/HD and pesticides w/HD.	ADL = (4 + 4.5 + 4.5 + 2 + 4.5)/5 = 3.9 (High ADL)
6	A/L <sub>2</sub>	28°40'29.61" S; 59°58'57.59" W	I <sub>1</sub> : Ground cover with implanted vegetation (pastures) and crop stubble. I <sub>2</sub> : Agriculture and livestock for more 25 years. I <sub>3</sub> : NT in the past 25 years plus MT (20 years). I <sub>4</sub> : 2 crops w/R. I <sub>5</sub> : organic fertilizer (cow manure) and agrochemicals (fertilizers w/RD and pesticides w/LD).	ADL = (3 + 3 + 4.5 + 3 + 2.5)/5 = 3.2 (Medium ADL)
7	A/L <sub>3</sub>	30°51'2.97" S; 61°10'46.72" W	I <sub>1</sub> : Implanted vegetation (pastures) and crop stubble. I <sub>2</sub> : Agriculture and livestock for 10 years (it was native woodland before). I <sub>3</sub> : NT less than 10 years. I <sub>4</sub> : 2 crops w/R. I <sub>5</sub> : Pesticides w/LD.	ADL = (3 + 3 + 1 + 3 + 2)/5 = 2.4 (Low ADL)
8	LW <sub>2</sub>	30°26'07.32" S; 60°28'22.38" W	I <sub>1</sub> : NVP with minimum extraction of trees. I <sub>2</sub> : ELNW for more than 20 years. I <sub>3</sub> : None. I <sub>4</sub> : Natural vegetation. I <sub>5</sub> : Only natural organic fertilizer (cattle dung) by livestock grazing.	ADL = (1 + 2 + 0 + 0 + 0.5)/5 = 0.7 (Low ADL)
9	A/L <sub>4</sub>	30°42'20.42" S; 60°46'14.01" W	I <sub>1</sub> : Implanted vegetation (pastures) and crop stubble. I <sub>2</sub> : Agriculture and livestock for more than 25 years. I <sub>3</sub> : NT in the past 25 years plus MT (20 years). I <sub>4</sub> : 2 crops w/R. I <sub>5</sub> : Chemical fertilizers w/HD and pesticides w/HD.	ADL = (3 + 3 + 4.5 + 3 + 4.5)/5 = 3.6 (High ADL)
10	A <sub>3</sub>	30°56'49.82" S; 60°40'31.43" W	I <sub>1</sub> : Minimum ground cover by crop stubble. I <sub>2</sub> : Intensive agriculture for 30 years. I <sub>3</sub> : NT (20 years) plus MT in last 10 years. I <sub>4</sub> : 1 crop without rotation (monoculture in the last 5 years). I <sub>5</sub> : Chemical fertilizers w/RD and pesticides w/HD.	ADL = (4 + 4.5 + 3 + 3.5 + 3)/5 = 3.6 (High ADL)
11	LW <sub>3</sub>	31°01'59.30" S; 60°05'29.94" W	I <sub>1</sub> : NVP with minimum extraction of trees. I <sub>2</sub> : ELNW for more 20 years. I <sub>3</sub> : None. I <sub>4</sub> : Natural vegetation. I <sub>5</sub> : Only natural organic fertilizer (cattle dung) by livestock grazing.	ADL = (1 + 2 + 0 + 0 + 0.5)/5 = 0.7 (Low ADL)
12	A/L <sub>5</sub>	31°04'02.54" S; 61°08'43.79" W	I <sub>1</sub> : Implanted vegetation (pastures) and crop stubble. I <sub>2</sub> : Agriculture and livestock for more than 20 years. I <sub>3</sub> : NT (20 years) more MT (in the last 10 years). I <sub>4</sub> : 2 crops w/R. I <sub>5</sub> : Chemical fertilizers w/RD and pesticides w/HD.	ADL = (3 + 3 + 3 + 3 + 3)/5 = 3 Medium ADL
13	IG	31°10'02.20" S; 60°13'42.94" W	I <sub>1</sub> : Loss of natural vegetation cover by animal overgrazing and trampling. I <sub>2</sub> : ELNG <sup>4</sup> for more than 25 years w/animal overload. I <sub>3</sub> : None. I <sub>4</sub> : Natural vegetation. I <sub>5</sub> : Pesticides w/LD.	ADL = (1 + 2*5 + 0 + 0 + 2)/5 = 2.6 (Medium ADL)
14	A/L <sub>6</sub>	31°15'34.70" S; 60°39'41.36" W	I <sub>1</sub> : Implanted vegetation (pastures) and crop stubble. I <sub>2</sub> : Agriculture and livestock for more 25 years. I <sub>3</sub> : NT (in the last 25 years) plus MT (20 years). I <sub>4</sub> : 2 crops w/R. I <sub>5</sub> : Chemical fertilizers w/HD and pesticides w/HD.	ADL = (3 + 3 + 4.5 + 3 + 4.5)/5 = 3.6 (High ADL)
15	A/L <sub>7</sub>	31°18'02.98" S; 61°02'14.40" W	I <sub>1</sub> : Implanted vegetation (pastures) and crop stubble. I <sub>2</sub> : Agriculture and livestock for 20 years. I <sub>3</sub> : NT (for 20 years) plus MT in last 10 years. I <sub>4</sub> : 3 crops w/R. I <sub>5</sub> : Organic fertilizer (cow manure) and agrochemicals (fertilizers w/RD and pesticides w/RD).	ADL = (3 + 3 + 3 + 2 + 2.5)/5 = 2.7 (Medium ADL)
16	H <sub>2</sub>	31°32'26.76" S; 0°41'28.66" W	I <sub>1</sub> : Year-round vegetation cover with season crops. I <sub>2</sub> : Intensive horticulture for more than 30 years. I <sub>3</sub> : NT (for 25 years) plus MT in last 20 years. I <sub>4</sub> : Polyculture (different crops) w/R. I <sub>5</sub> : Organic fertilizer (cow and chicken manure) and agrochemicals (fertilizers w/RD and pesticides w/LD).	ADL = (2 + 3.5 + 4.5 + 1 + 2.5)/5 = 2.9 (Medium ADL)
17	H <sub>3</sub>	31°34'47.00" S; 60°41'4.00" W	I <sub>1</sub> : Implanted vegetation (vegetable crops) and crop stubble. I <sub>2</sub> : Intensive horticulture for more than 30 years. I <sub>3</sub> : NT for 25 years plus MT in last 20 years. I <sub>4</sub> : 2 crops w/R. I <sub>5</sub> : Chemical fertilizers w/HD and pesticides w/HD.	ADL = (3 + 3.5 + 4.5 + 3 + 4.5)/5 = 3.7 (High ADL)
18	A <sub>4</sub>	31°36'45.78" S; 61°05'33.60" W	I <sub>1</sub> : Minimum ground cover by crop stubble. I <sub>2</sub> : Intensive agriculture for more 25 years. I <sub>3</sub> : NT for 25 years plus MT in last 18 years. I <sub>4</sub> : 3 crops w/R. I <sub>5</sub> : Chemical fertilizers w/RD and pesticides w/HD.	ADL = (4 + 4.5 + 4.5 + 2 + 3)/5 = 3.6 (High ADL)
19	A <sub>5</sub>	33°02'23.79" S; 60°52'57.25" W	I <sub>1</sub> : Minimum ground cover by crop stubble. I <sub>2</sub> : Intensive agriculture for 25 years. I <sub>3</sub> : NT for 25 years plus MT for 20 years. I <sub>4</sub> : 2 crops w/R. I <sub>5</sub> : Chemical fertilizers w/RD and pesticides w/HD.	ADL = (4 + 4.5 + 4.5 + 3 + 3)/5 = 3.8 (High ADL)
20	A <sub>6</sub>	34°12'41.8" S; 62°37'01.0" W	I <sub>1</sub> : Minimum ground cover by crop stubble. I <sub>2</sub> : Intensive agriculture for 10 years (before was native woodland). I <sub>3</sub> : NT for 10 years. I <sub>4</sub> : 3 crops w/R. I <sub>5</sub> : Chemical fertilizers w/RD and pesticides w/HD.	ADL = (4 + 4.5 + 1 + 2 + 3)/5 = 2.9 (Medium ADL)

References: LW: Livestock in Woodland, IG: Livestock in Grassland, A: Agriculture/Livestock, H: Horticulture. The sites analyzed are renamed according to their use classification and latitudinal order.  
<sup>1</sup> NVP: Natural Vegetation Preserved. <sup>2</sup> ELNW: Extensive Livestock in Native Woodland. <sup>3</sup> R: Rotation. <sup>4</sup> ELNG: Extensive Livestock in Natural Grassland.

(1934); pH, by potentiometric method in a soil–water ratio of 1:2.5 w/v (Jackson, 1976); BD, by cylinder method (Baver et al., 1973); CEC, using 1 N pH 7 ammonium acetate for the extraction of interchangeable cations, such as Calcium (Ca<sup>++</sup>) and Magnesium (Mg<sup>++</sup>) by complexometry, and Sodium (Na<sup>+</sup>) and Potassium (K<sup>+</sup>) by flame photometry (MAG, 1982); and soil texture, by pipetting method for granulometric analysis of fractions smaller than 62 μm (Gee and Bauder, 1986). Water retention ability and nutrients availability are related to soil texture (percentage of clay and sand).

By randomly subsampling (or subgrouping) the 40 soil monoliths of each site and estimating the averages of the variables  $D_T$ ,  $H$ , %Ad, and  $D_3$  for the monoliths selected, it was determined that these averages do not change significantly ( $p < 0.01$ ) for sample sizes greater than 15 monoliths per site. Therefore, fewer monoliths could have been used.

### 2.2.3. Backgrounds and surveys for the sampled agroecosystems

The sampled sites correspond to agroecosystems with different land use and management backgrounds (Masin, 2017). The ADL score (dimensionless) of each site analyzed is empirically obtained by using Table 1 as a survey from information on land use history and management practices. The conditions registered for each specific indicator and the ADL score obtained for each surveyed site are summarized in Table 2. Regarding the use of agrochemicals, classified as LD, RD or HD in Table 1, the following was inquired: types, brands, concentrations and number of applications per year. The agrochemical types, concentrations and number of applications are described in Masin (2017).

### 2.3. Statistical modeling and development of the ADL estimator

The following four steps must be performed to develop the ADL estimator: the first step is to validate that the proposed mapping (Table 1) results in an empirical assessment (or score) correlated with the biological data that characterize the earthworm community sampled. The second step is to select the most relevant variables to the multivariate correlation model, analyzing the correlations of the edaphic and biological variables with the ADL score. Subsequently, considering only the subset of variables selected, the third step is to calibrate a multivariate correlation model by using 10-fold cross-validation and its consistency must be verified (Martens and Naes, 1989). The Principal Component Analysis (PCA) allows exploring the correlations between these variables in a reduced space and provides a framework for understanding the displacements of the latent variables in relation to the original ones and the correlations between the original variables (Godoy et al., 2014). In order to adjust a correlation model between the selected variables, the observations of the vector  $\mathbf{z}$  (centered and scaled) are entered in the following data matrix ( $20 \times$  selected variables):

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}'_1 \\ \vdots \\ \mathbf{z}'_{20} \end{bmatrix} \quad (2)$$

where  $\mathbf{z}'_i$  is the  $i$ -th multivariate observation collected with  $i = 1, \dots, 20$ . Then, PCA modeling fits a multivariate correlation matrix  $(n - 1)^{-1}\mathbf{Z}'\mathbf{Z} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$  (see Appendix A) to project observations of possibly correlated variables  $\mathbf{z}$  to a latent space where the latent variables  $\mathbf{t}$  are linearly uncorrelated as follows:  $\mathbf{t} = \mathbf{P}'\mathbf{z}$ , where  $\mathbf{P}$  is the eigenvectors matrix that diagonalizes the correlation matrix (Godoy et al., 2014).

Finally (fourth step), the ADL estimator is constructed based on this correlation model. In PCA, the vector  $\mathbf{z}$  can be divided into a known part  $\mathbf{x}$  and an unknown part  $\mathbf{y}$ . Consequently, the adjusted latent model  $\mathbf{P}$  can be partitioned as follows:

$$\mathbf{z} = \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_y \\ \mathbf{P}_x \end{bmatrix} \mathbf{t} \quad (3)$$

where  $\mathbf{t}$  is the latent vector corresponding to  $\mathbf{z}$  projection (i.e.,  $\mathbf{t} = \mathbf{P}'\mathbf{z}$ ). Then, when only  $\mathbf{x}$  (the known part of  $\mathbf{z}$ ) is used to estimate the latent vector  $\mathbf{t}$ , the following relationship is obtained from the bottom of Eq. (3):

$$\hat{\mathbf{t}} = (\mathbf{P}'_x\mathbf{P}_x)^{-1}\mathbf{P}'_x\mathbf{x} \quad (4)$$

Therefore,  $\mathbf{y}$  can be predicted from  $\mathbf{x}$  by using the top of Eqs. (3) and (4) as follows:

$$\hat{\mathbf{y}} = \mathbf{P}_y(\mathbf{P}'_x\mathbf{P}_x)^{-1}\mathbf{P}'_x\mathbf{x} \quad (5)$$

Finally, the predicted value at original units is obtained by re-scaling and re-centering Eq. (5) as follows:

$$\widehat{ADL} = D_y[\mathbf{P}_y(\mathbf{P}'_x\mathbf{P}_x)^{-1}\mathbf{P}'_x]\mathbf{D}_x^{-1}(\mathbf{x}_0 - \bar{\mathbf{x}}) + \bar{y} \quad (6)$$

where  $\mathbf{x}_0$  is the measurements vector  $\mathbf{x}$  at original units,  $\bar{\mathbf{x}}$  is the mean vector, and  $\mathbf{D}_x$  is a scaling matrix with standard deviations on its diagonal. This equation represents the ADL estimator for sites with agricultural and livestock use in Santa Fe province.

## 3. Results and discussions

The biological and edaphic data of each sampled site are summarized in Table 3. None of the sampled sites has a BD that could affect biological data and, consequently, this variable is excluded from the data analyzed (BD < 1.38 [g/cm<sup>3</sup>]) (Domínguez et al., 2009).

### 3.1. Relationship between ADL and biological data

The values of total density (and percentage of adults) of each site (Table 3) are grouped in three subsets (low, medium and high ADL subset) according to their ADL scores (Table 2). Then, the earthworm populations (characterized by total density and percentage of adults) in each site varies significantly ( $p < 0.05$  for the Kruskal-Wallis test) for different ADL values ordered in low, medium and high ADL subset (i.e., it varies for different types of land use and management practices). Regarding tillage, all the sites show percentages of juveniles notably higher than those of adults (Table 3; Fig. 2). In particular, high ADL sites have the highest percentages of juveniles (between 80 and 100%). This is consistent with Jiménez et al. (2012), Lemtiri et al. (2014) and Bertrand et al. (2015), who reported that a high proportion of juveniles is frequent in soils intensively used, where land management practices such as conventional tillage, monoculture, and high doses of agrochemicals (especially pesticides) impact on the decrease in biological activity, on the growth retardation and on the maturation of the individuals.

Thirteen earthworm species were found in the 20 sites sampled, and were grouped into five families: Acanthodrilidae (*Dichogaster bolau*, *Microscolex dubius*), Glossoscolecidae (*Glossodrilus parecis*), Lumbricidae (*Aporrectodea caliginosa*, *A. rosea*, *A. trapezoides*, *Bimastos parvus*, *Eisenia fetida*, *Octolasion tyrtaeum*), Megascolecidae (*Amyntas morrisi*, *Metaphire californica*) and Ocerodrilidae (*Eukerria saltensis*, *E. stagnalis*). Lumbricidae family represented 46% of the species found, where four are native (*M. dubius*, *G. parecis*, *E. saltensis*, *E. stagnalis*) and the other ones are exotic. Earthworm species richness (S) were ordered according to the ADL value of each site (Fig. 3). Low and medium ADL sites exhibited earthworm assemblages with species richness higher than that of high ADL sites (75% of these sites presented a maximum of two species).

At low ADL sites, 11 species were found: *D. bolau*, *M. dubius*, *A. trapezoides*, *A. rosea*, *O. tyrtaeum*, *A. morrisi*, *M. californica*, *E. stagnalis*, *E. saltensis*, *A. caliginosa* and *E. fetida*; while medium ADL sites exhibited 10 species, including the first eight species of the previous group (i.e., 8 species in common) plus the following 2 species: *B. parvus* and *G. parecis*. Finally, at high ADL sites, the following 7 species were found: *M. dubius*, *A. rosea*, *A. trapezoides*, *O. tyrtaeum*, *A. morrisi*, *M. californica* and

**Table 3**  
Biological and edaphic data for the sites sampled in Santa Fe province.

Latitudinal order	Site-land use type	Earthworm mean density [ind./m <sup>2</sup> ]		Density of each category [ind./m <sup>2</sup> ]		H	SOM [%]	pH	BD [g/cm <sup>3</sup> ]	CEC [meq.100 g]			Soil texture [%]		
		D <sub>T</sub>	Adult density/% Ad	D <sub>1</sub> + D <sub>2</sub>	D <sub>3</sub>					Ca <sup>++</sup>	Mg <sup>++</sup>	K <sup>+</sup>	Sand	Silt	Clay
1	LW <sub>1</sub>	355	47/13	0	355	1.09	7.90	6.60	0.98	27.60	6.60	4.80	11.80	48.50	39.70
2	A <sub>1</sub>	119	0/0	0	119	0.34	1.87	5.80	1.03	11.00	0.40	0.44	27.70	46.20	26.40
3	A/L <sub>1</sub>	102	26/25	0	102	1.28	1.56	5.80	1.07	7.80	0.40	0.56	29.1	39.5	21.5
4	H <sub>1</sub>	560	94/17	264	296	1.48	5.17	6.80	1.24	24.80	4.6	0.36	41	36.8	22.2
5	A <sub>2</sub>	107	23/21	27	80	1.08	1.85	6.30	1.06	5.20	4.80	1.80	22.20	54.80	23.00
6	A/L <sub>2</sub>	143	17/12	91	52	1.36	3.20	6.60	0.91	12.80	1.60	1.78	12.10	53.30	34.70
7	A/L <sub>3</sub>	203	16/8	1	202	0.53	4.35	7.00	1.05	24.00	0.40	1.90	3.60	63.20	33.30
8	LW <sub>2</sub>	140	43/31	51	89	1.59	7.15	6.50	0.83	15.20	0.6	1.92	7.8	60.7	31.5
9	A/L <sub>4</sub>	42	0/0	0	42	0.49	2.56	6.20	1.16	6.60	6.00	1.80	23.70	56.10	20.20
10	A <sub>3</sub>	16	0/0	6	10	0.66	2.76	6.10	1.12	10.80	3.00	1.30	24.60	68.70	6.70
11	LW <sub>3</sub>	92	25/27	17	75	1.22	5.05	6.30	1.04	19.40	2.2	0.28	26	50	23
12	A/L <sub>5</sub>	18	0/0	0	18	0.62	3.29	5.90	1.26	12.80	0.40	1.80	5.90	60.80	33.30
13	LG	110	31/28	8	102	0.69	2.40	7.00	1.35	21.00	0.40	0.10	63.20	23.10	13.70
14	A/L <sub>6</sub>	246	2/1	97	149	0.67	2.16	6.20	1.35	5.76	2.44	1.44	12.10	61.60	26.30
15	A/L <sub>7</sub>	100	6/6	6	94	0.33	2.96	6.60	1.30	16.00	0.20	1.60	27.20	61.30	11.50
16	H <sub>2</sub>	95	35/37	21	74	1.40	2.60	6.50	1.22	9.80	0.80	1.20	20.10	60.30	19.60
17	H <sub>3</sub>	17	0/0	0	17	0	1.53	7.20	1.05	7.20	2.80	1.44	23.60	58.30	18.10
18	A <sub>4</sub>	71	23/32	3	68	0.69	3.42	5.60	1.09	12.20	1.00	1.40	4.90	65.90	29.20
19	A <sub>5</sub>	29	1/3	0	29	0.29	2.18	6.10	1.11	10.20	1.10	1.12	12.70	26.10	61.20
20	A <sub>6</sub>	79	13/16	3	76	1.42	2.18	5.60	1.21	9.60	0.20	1.00	13.80	14.60	71.60

References: LW: Livestock in Woodland, LG: Livestock in Grassland, A: Agriculture, A/L: Agriculture/Livestock, H: Horticulture.

*E. stagnalis*, which can be also found at low and medium ADL sites. The three ecological categories (epigeic, epiendogeic and endogeic) were present only at H<sub>1</sub>, LW<sub>3</sub> and H<sub>2</sub> with high species richness (Fig. 3). The presence of species of different ecological categories is a consequence of the spatial heterogeneity, the availability of nutrients and the application of conservation tillage practices (Brown et al., 2004; Pelosi et al., 2009). At the other low and medium ADL sites, the assemblies were constituted by two categories (epiendogeic and endogeic). Seventy-five percent of the high ADL sites presented assemblages with epiendogeic and endogeic species, but with lower species richness than that of the low and medium ADL sites, and the remaining 25% exhibited assemblages constituted only by endogeic species (Fig. 3). Intensive agricultural practices modify the physical and edaphic environment and the availability of resources, which affects the presence and activities of earthworm species, epigeics being the most negatively affected with respect to endogeics (Bartz et al., 2014). In this sense, endogeic

earthworms of the genus *Aporrectodea* (*A. trapezoides* and *A. rosea*) were the predominant species for most of the sites sampled, and *A. morrisi* (epiendogeic species) was recorded in sites with high ADL (A/L<sub>4</sub>, A<sub>3</sub>, H<sub>3</sub>, A<sub>2</sub>) and low SOM. This exotic species is frequent and it is becoming dominant in agroecosystems with intensive soil use; due to its greater tolerance to anthropic disturbances (Brown et al., 2004; Chan and Barchia, 2007). On the contrary, the native species *E. stagnalis* was found mostly at sites with low and medium ADL (H<sub>1</sub>, LW<sub>1</sub>, LW<sub>2</sub>, LW<sub>3</sub>, and H<sub>2</sub>) and with high moisture or close to bodies of water. This native species is an indicator of high moisture and acidity in soil with good to deteriorated quality (Falco et al., 2015).

In summary, the empirical evaluation proposed (ADL) is consistent with the biological data registered.

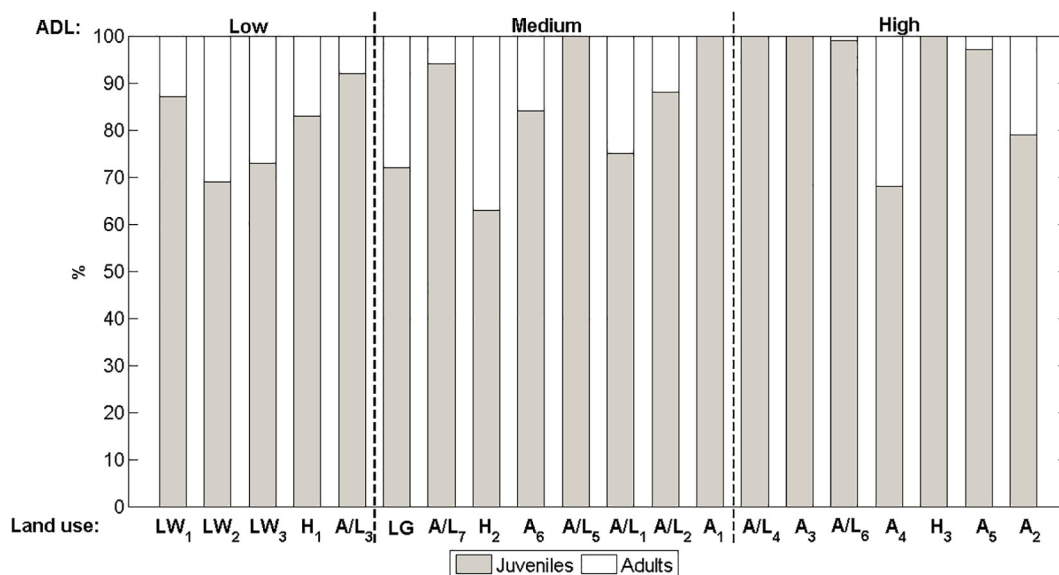


Fig. 2. Percentages of adults and juveniles in different sites ordered according to their ADL value.

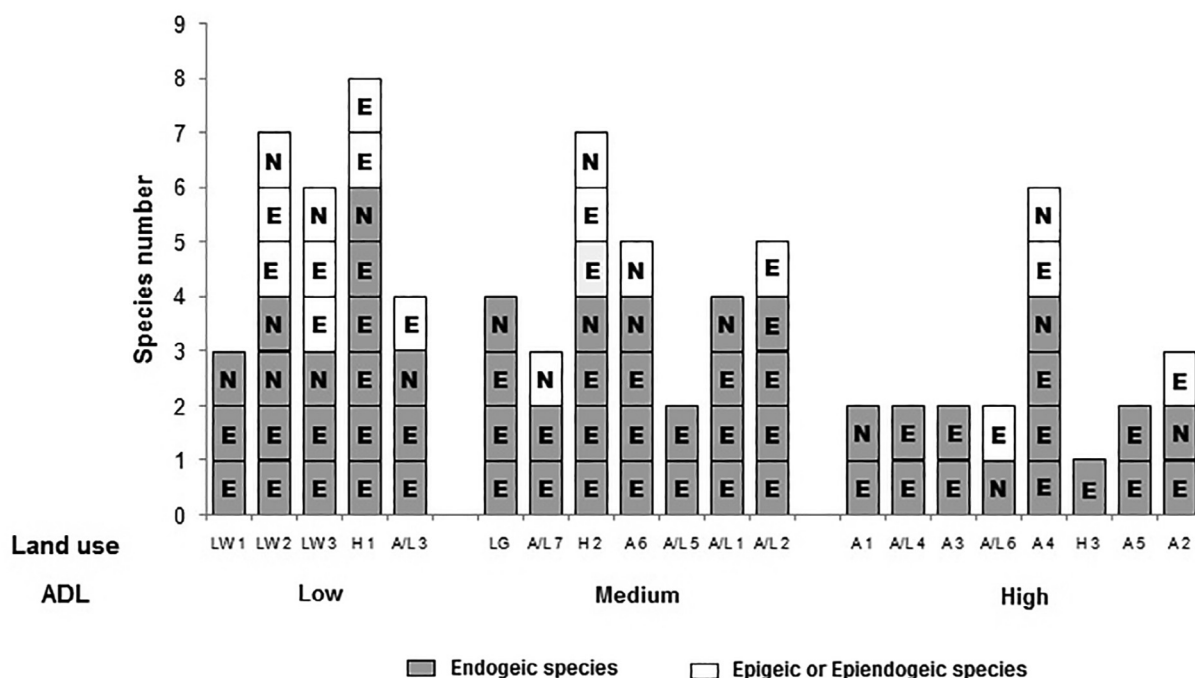


Fig. 3. Earthworm species richness in sites with low, medium and high ADL. References: N: Native species; E: Exotic species.

### 3.2. Correlation structure between earthworm populations, edaphic properties and ADL

In this section, the correlations between earthworm population characteristics, soil physical–chemical properties, and the ADL for each site are analyzed. The number of samples collected (20) is low in relationship to the number of edaphic properties and population characteristics (Table 3). Therefore, fewer variables than those available should be selected in order to reliably adjust a multivariable correlation structure (Martens and Naes, 1989). Based on the correlation coefficient of ADL with other variables and the cross-correlations, the following variables were selected (see Fig. A.1 in Appendix A): edaphic data (SOM, pH, and Ca<sup>++</sup>), biological data (%Ad, D<sub>1</sub> + D<sub>2</sub>, D<sub>3</sub>, H), and ADL. Notice that Clay and Mg<sup>++</sup> have low correlation coefficients (*r* < 0.1, considered threshold) and that D<sub>T</sub> is highly correlated (*r* greater than 0.6) with D<sub>1</sub> + D<sub>2</sub> and D<sub>3</sub>; hence, these variables were not selected (see Fig. A.1 in Appendix A).

Earthworm populations are characterized by the following variables: D<sub>1</sub> + D<sub>2</sub>, D<sub>3</sub>, %Ad, H; and edaphic properties are represented by: SOM, pH, Ca<sup>++</sup>. Then, the measurement vector *z* (8x1) is composed as follows:

$$z' = [ADL, (D_1 + D_2), D_3, \%Ad, H, SOM, pH, Ca^{++}] \quad (7)$$

PCA is performed to determine if there exists a relationship between earthworm population, edaphic properties, and the proposed empirical assessment of ADL. PCA involves a decomposition of the dataset *Z* (Eq. (2)) along the directions of maximum variability. The 95% of the variability of the data analyzed can be represented by retaining 5 PCs in the model, which highlights that there are 3 collinearities among the 8 variables considered. This result demonstrates that there exist cause-effect relationships between the variables considered, which validates the initial hypothesis of modeling: adding the ADL to edaphic properties allows describing the incident effect on earthworm populations.

Two PCA biplots considering only the two first principal components (PC<sub>1</sub>, PC<sub>2</sub>) and (PC<sub>1</sub>, PC<sub>3</sub>) are shown in Fig. 4a and 4b, respectively. The percentage of variance explained by each PC is shown on its respective axis. The dispersion of the projected observations on this reduced space represents approximately the 80% of the data variability. These biplots allow visualizing the magnitude and sign of the

contribution of each original variable (ADL, D<sub>1</sub> + D<sub>2</sub>, D<sub>3</sub>, %Ad, H, SOM, pH, Ca<sup>++</sup>) to the first three principal components (PC<sub>1</sub>, PC<sub>2</sub> and PC<sub>3</sub>) and how each site (observation *z*) is represented in terms of their latent coordinates (Eq. (3)). As it can be seen from the relationships between the biological and edaphic data and ADL, low and medium ADL sites have positive projections on D<sub>1</sub> + D<sub>2</sub>, D<sub>3</sub>, %Ad, and H, unlike high ADL sites, which have negative projections (Fig. 4). The value of the correlation between two original variables is given by the projection of a variable contribution on the axis of the other contribution. The correlation structure (or correlations between the variables) obtained is consistent, as D<sub>1</sub> + D<sub>2</sub>, D<sub>3</sub>, %Ad, H, pH, Ca<sup>++</sup> and SOM have negative correlations with ADL (Fig. 4). In general, ADL is in the opposite direction to biological data, thus confirming that the proposed assessment is a reliable indicator of soil quality degradation affecting negatively earthworm populations. Another result that confirms the modeling hypothesis is: if the variable ADL is excluded from *z* Eq. (7) and a PCA model is fitted to these data, then the correlations obtained are not consistent.

### 3.3. Inferential estimator of the ADL in agroecosystems

In this section, the use of earthworms as a bioindicator sensitive to edaphic properties and anthropic disturbances is proposed to estimate the ADL of agricultural production systems in Santa Fe province (Argentina). Thus, the ADL of an agroecosystem could be estimated by counting the earthworm species and measuring the edaphic properties of a given soil sample. The prediction error for each site (Fig. 5a) is computed as the ADL observed by using the survey of Table 1 (i.e., the true ADL score) minus the ADL estimated by using Eq. (6) (i.e., prediction error = true ADL score – estimated ADL score). Prediction errors are lower than 0.6 (in absolute value) for all the sites, except for sites LW<sub>3</sub> and A<sub>4</sub> (Fig. 5a).

The ADL estimator is also calibrated by Partial Least Squares (PLS) method (Godoy et al., 2014) and the prediction errors are lower than those for the PCA-based estimator, except for sites LW<sub>3</sub> and A<sub>4</sub>, where the errors are higher (Fig. 5a and b). This confirms, once more, that these sites are outliers of the model. In particular, site LW<sub>3</sub> has biological and edaphic data with low values, which implies that the ADL estimator produces a predicted value higher than the observed value.

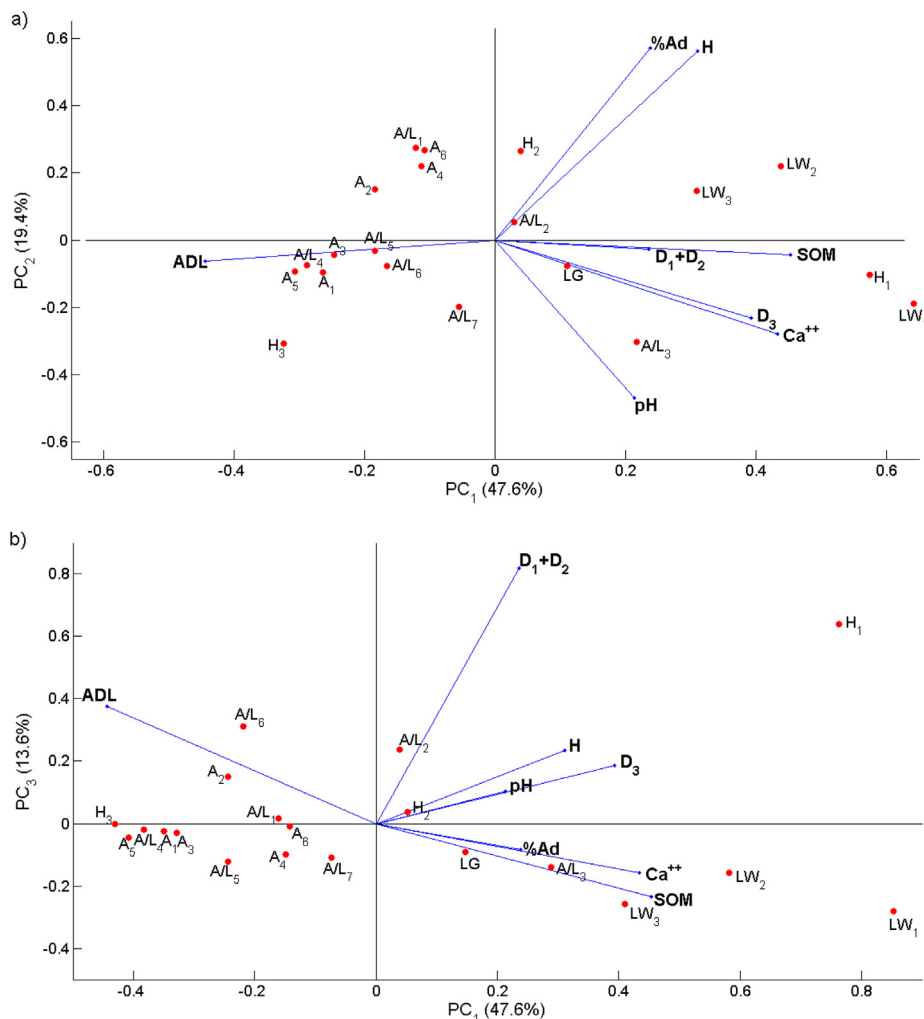


Fig. 4. Correlation structure in the reduced space, where 3 PCs explain the 80% of the data variability. a) Biplot based on the two first principal components of data matrix (PC<sub>1</sub> and PC<sub>2</sub>). b) Biplot based on the first principal component (PC<sub>1</sub>) and third principal component (PC<sub>3</sub>).

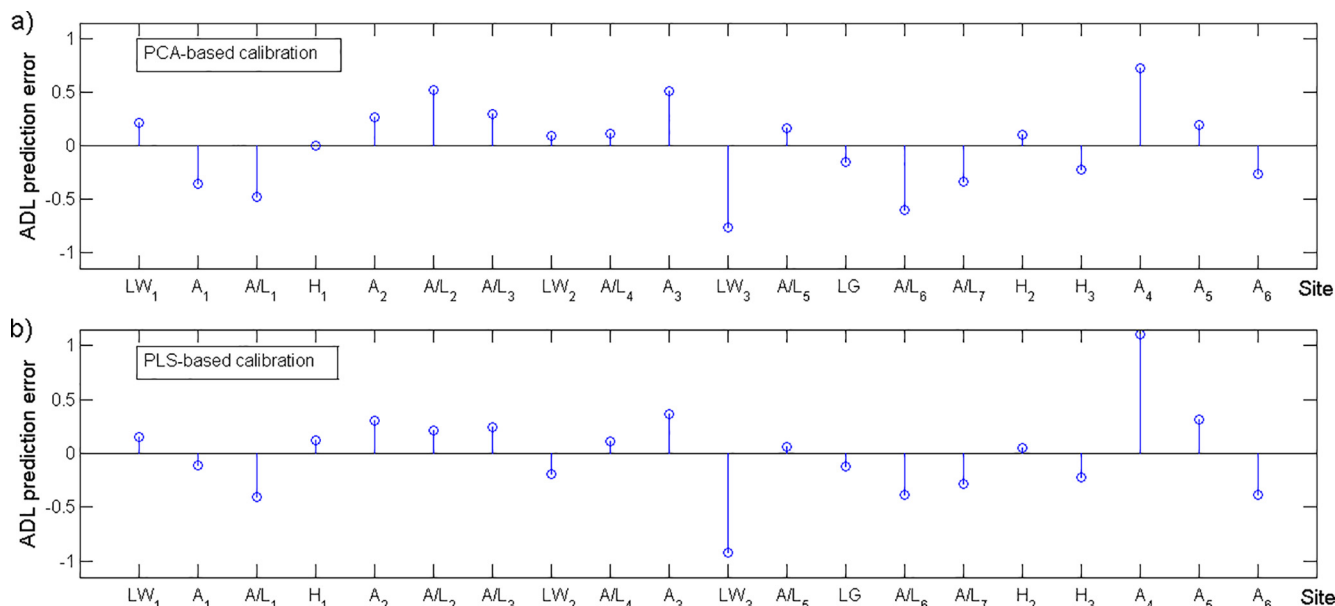


Fig. 5. ADL prediction error for each site (computed as the ADL observed by using the survey of Table 1 minus the ADL estimated by using biological and edaphic data): a) For the model calibrated with PCA method. b) For the model calibrated with PLS method.



The earthworm population density of each soil sample is related to its water content (Johnston et al., 2018); however, this edaphic property was not measured. Probably, the accuracy of the ADL estimator can be improved by adding this measurement to the model proposed. The prediction for site A<sub>4</sub> has an overestimated value. This significant error can be due to a mistake in the background-survey registered, to the fact that the proposed survey is incomplete or to the fact that the adjusted model does not consider this case. This indicates that the observability of the proposed ADL estimator can be improved by adding another specific indicator to the survey.

### 3.4. Example of the use for monitoring

To avoid the degradation of soils, fertility and sustainability must be preserved. It can be achieved by applying rationally technical criteria, crop rotation and diversity, replenishment of nutrients, and the set of good agricultural practices. In consequence, it is necessary to evaluate the impact of land management practices on the health of agroecosystems.

When the ADL estimator is used to monitor agroecosystems, it is useful to diagnose the root cause of the predicted value. For example, for site A<sub>2</sub> the predicted ADL is 3.65 (ADL observed is 3.9 in Table 2) and the specific indicators I<sub>1</sub>, I<sub>2</sub> and I<sub>3</sub> can be determined by inspection or by tax information as follows: I<sub>1</sub> = 4, I<sub>2</sub> = 4.5, I<sub>3</sub> = 4.5. Then, from the formula of ADL (Eq. (1)) it is obtained that: I<sub>4</sub> + I<sub>5</sub> = 5.25, hence, the mean value inferred for I<sub>4</sub> and I<sub>5</sub> is 2.625 (5.25/2). From Table 1, the following possibilities result for these two specific indicators (i.e. with a total close to 5.25): I<sub>4</sub> + I<sub>5</sub> = 2 + 3, = 3 + 2.5, or = 3.5 + 2. However, the true combination is I<sub>4</sub> + I<sub>5</sub> = 2 + 4.5 = 6.5 (producing mean values inferred for I<sub>4</sub> and I<sub>5</sub> equal to 3.25) but this possibility was not among the previous ones considered. This is because the ADL prediction error 0.25 is propagated to the specific indicators I<sub>4</sub> + I<sub>5</sub> with a value of 1.25, which made it difficult to diagnose the true value of these two indicators.

## 4. Conclusions

There are several studies linking soil quality with biological and edaphic data. However, to our knowledge, the presented ADL estimator is the first one to assess the health degradation of an agroecosystem by integrating biological data, edaphic data, and management practices.

The proposed assessment evaluates the disturbance level (which is

## Appendix A

### Decomposition on principal components

The PCA modeling decomposes  $\mathbf{Z}$  into score vectors ( $\mathbf{t}_a$ ), loadings vectors ( $\mathbf{p}_a$ ) and residual errors ( $\tilde{\mathbf{Z}}$ ) as follows (Godoy et al., 2014):

$$\mathbf{Z} = \sum_{a=1}^A \mathbf{t}_a \mathbf{p}'_a + \tilde{\mathbf{Z}} = \mathbf{TP}' + \tilde{\mathbf{Z}}, \quad \mathbf{T} = [\mathbf{t}_1 \dots \mathbf{t}_A], \quad \mathbf{P} = [\mathbf{p}_1 \dots \mathbf{p}_A] \quad (\text{A.1})$$

where  $A$  is the number of principal components (PCs) retained in the model. The scores  $\mathbf{T}$  (or observations of the latent variables) can be represented in terms of the original data  $\mathbf{Z}$  as follows:

$$\mathbf{T} = \mathbf{ZP} \quad (\text{A.2})$$

since  $\mathbf{P}'\mathbf{P} = \mathbf{I}$  and  $\tilde{\mathbf{Z}}\mathbf{P} = 0$ . The matrix  $\mathbf{P}$  unambiguously defines the decomposition of  $\mathbf{Z}$  as follows:  $\mathbf{Z}$  is projected to the latent space through  $\mathbf{P}$  (Eq. (A.2)), and it is reconstructed by means of  $\mathbf{P}'$  (Eq. (A.1)). PCA allows estimating the multivariate correlation matrix as follows:

$$(n-1)^{-1} \mathbf{Z}'\mathbf{Z} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' \quad (\text{A.3})$$

where  $\mathbf{P}$  is the eigenvectors matrix associated with the nonzero eigenvalues matrix  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_A) = (n-1)^{-1} \mathbf{T}'\mathbf{T}$ .

### Selection of the most relevant variables

The correlations between the ADL score (i.e., the artificial variable) and the edaphic and biological variables are shown in Fig. A1 together with the Pearson's correlation coefficients.

related to the health degradation of an agroecosystem) instead of evaluating the health of the agroecosystem (which is related to soil quality). This is an advantage, since it needs no references because the ADL score increases only when the agroecosystem is disturbed.

ADL estimator proved to be a valuable methodology to monitor soil health degradation in different agroecosystems in Santa Fe province (Argentina). As it is based on a data-driven model which is easy to apply, it could be also used to monitor the sustainability of agroecosystems in other regions of the world. It is a remarkable benefit since the monitoring of agroecosystem health allows taking actions to improve the sustainability of the ecosystem services.

This work shows that by using a mapping of land use and management practice characteristics, the quantification of land degradation due to anthropic factors can be significantly improved. In addition, it is shown that earthworm populations (characterized by species density, diversity, and activity) are negatively according to the intensity of land use and management practices (i.e. the ADL).

To increase the accuracy of the ADL estimator proposed, it is necessary to use more samples (or sites) from the geographic area studied and to improve the mapping of land use and management practice characteristics by using fuzzy sets or neural networks, or any other machine learning technique. However, this work aims to be a first approach to how to evaluate ADL in agroecosystems.

## CRedit authorship contribution statement

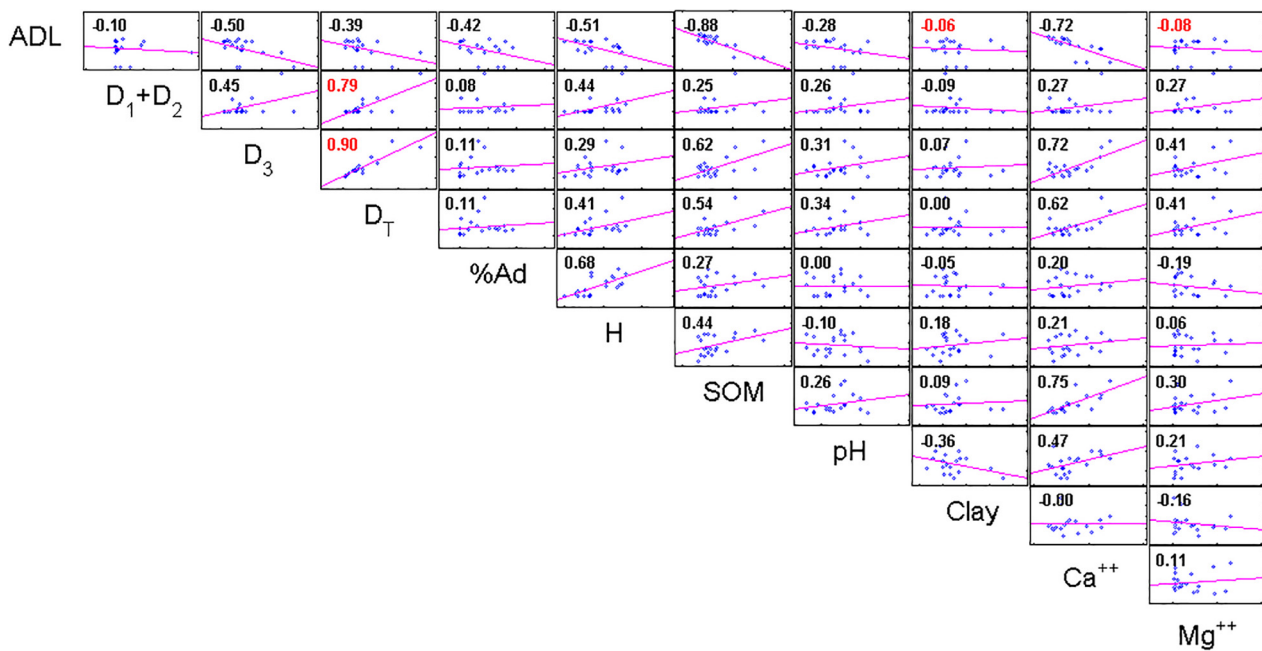
C. Masin: Conceptualization, Investigation, Methodology, Validation. A.R. Rodríguez: Investigation, Validation. C. Zalazar: Resources, Supervision. J.L. Godoy: Conceptualization, Methodology, Software, Formal analysis, Visualization.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors are grateful for the financial support received from CONICET, and ANPCyT (Argentina).



**Fig. A1.** Correlations between the variables (together with correlation coefficient). Anthropogenic disturbance level (ADL), Population total density,  $D_T$  [individuals/m<sup>2</sup>], percentage of adults (%Ad), species diversity ( $H$ ), Soil organic matter (SOM), Density for each category: Epigeic ( $D_1$ ), Epiendogeic ( $D_2$ ) and Endogeic ( $D_3$ ), where  $D_T = D_1 + D_2 + D_3$ .

## References

- Álvarez, C.R., Fernández, P.L., Taboada, M.A., 2012. Relación de la inestabilidad estructural con el manejo y propiedades de los suelos en la región Pampeana. *Ciencia del Suelo* 30 (2), 173–178.
- Bartz, M.L.C., Pasini, A., Brown, G.G., 2013. Earthworms as soil quality indicators in Brazilian no-tillage systems. *Appl. Soil Ecol.* 69, 39–48.
- Bartz, M.L.C., Brown, G.G., da Rosa, M.G., Klauberg Filho, O., James, S.W., Decaens, T., Baretta, D., 2014. Earthworm richness in land-use systems in Santa Catarina. *Brazil. Appl. Soil Ecol.* 83, 59–70.
- Baver, L.D., Gardner, W.H., Gardner, W.R., 1973. *Física de suelos*. Grupo Noriega Ed. UTEHA. México. p. 529.
- Bedano, J.C., Domínguez, A., 2016. Large-scale agricultural management and soil meso and macrofauna conservation in the Argentine Pampas. *Sustainability* 8 (7), 653.
- Bertrand, M., Barot, S., Blouin, M., Whalen, J., de Oliveira, T., Roger-Estrade, J., 2015. Earthworm services for cropping systems. *Agron. Sustain. Dev.* 35 (2), 553–567.
- Biasatti, L.V.N.R., Rozzatti, J.C., Fandiño, B., Pautaso, A., Mosso, E., Marteleur, G., Algañaz, N., Giraudo, A., Chiarulli, C., Romano, M., Ramírez Llorens, P., 2016. Las ecoregiones, su conservación y las áreas naturales protegidas de la provincia de Santa Fe. *Ministerio de Medio Ambiente, Santa Fe*, p. 244. [https://www.santafe.gov.ar/index.php/web/content/download/229660/1202209/file/LIBRO%20ECOREGIONES\\_web.pdf](https://www.santafe.gov.ar/index.php/web/content/download/229660/1202209/file/LIBRO%20ECOREGIONES_web.pdf).
- Blouin, M., Hodson, M.E., Delgado, E.A., Baker, G., Brussaard, L., Butt, K.R., Dai, J., Dendooven, L., Peres, G., Tondoh, J.E., Cluzeau, D., Brun, J.J., 2013. A review of earthworm impact on soil function and ecosystem services. *Eur. J. Soil Sci.* 64 (2), 161–182.
- Bouché, M.B., 1977. Strategies lombriciennes. *Ecol. Bull.* 25, 122–132.
- Brown, G.G., Moreno, A.G., Barois, I., Fragoso, C., Rojas, P., Hernandez, B., Patrón, J.C., 2004. Soil macrofauna in SE Mexican pastures and the effect of conversion from native to introduced pastures. *Agric. Ecosyst. Environ.* 103 (2), 313–327.
- Burkart, R., Barbaro, N.O., Sánchez, R.O., Gómez, D.A., 1999. Eco-regiones de la Argentina. Programa de desarrollo institucional, componente de política ambiental, Administración de Parques Nacionales. Buenos Aires Argentina 42.
- Butt, K.R., Lowe, C.N., 2004. Anthropogenic influences on earthworm distribution, Isle of Rum National Nature Reserve, Scotland. *Eur. J. Soil Biol.* 40 (2), 63–72.
- Carreño, L.V., Viglizzo, E.F., 2010. Efecto de la agricultura sobre la provisión de servicios ecosistémicos. Expansión de la frontera agropecuaria en Argentina y su Impacto ecológico-ambiental. INTA, Buenos Aires, Argentina. pp. 47–53.
- Carnovale, D., Baker, G., Bissett, A., Thrall, P., 2015. Earthworm composition, diversity and biomass under three land use systems in south-eastern Australia. *Appl. Soil Ecol.* 88, 32–40.
- Chan, K.Y., Barchia, I., 2007. Soil compaction controls the abundance, biomass and distribution of earthworms in a single dairy farm in south-eastern Australia. *Soil Tillage Res.* 94 (1), 75–82.
- Culman, S.W., Young-Mathews, A., Hollander, A.D., Ferris, H., Sánchez-Moreno, S., O'Geen, A.T., Jackson, L.E., 2010. Biodiversity is associated with indicators of soil ecosystem functions over a landscape gradient of agricultural intensification. *Landsc. Ecol.* 25, 1333–1348.
- Cunha, L., Brown, G.G., Stanton, D.W., Da Silva, E., Hansel, F.A., Jorge, G., McKey, D., Vidal-Torrado, P., Macedo, R.S., Velásquez, E., James, S.W., Lavelle, P., Kille, P., 2016. Soil animals and pedogenesis: the role of earthworms in anthropogenic soils. *Soil Sci.* 181 (3/4), 110–125.
- Dewi, W.S., Senge, M., 2015. Earthworm diversity and ecosystem services under threat. *Rev. Agr. Sci.* 3, 25–35.
- Domínguez, A., Bedano, J.C., Becker, A.R., 2009. Cambios en la comunidad de lombrices de tierra (Annelida: Lumbricina) como consecuencia del uso de la técnica de siembra directa en el centro-sur de Córdoba, Argentina. *Ciencia del Suelo* 27 (1), 11–19.
- Domínguez, A., Jiménez, J.J., Ortíz, C.E., Bedano, J.C., 2018. Soil macrofauna diversity as a key element for building sustainable agriculture in Argentine Pampas. *Acta Oecol. – Int. J. Ecol.* 92, 102–116.
- Ernst, G., Emmerling, C., 2009. Impact of five different tillage systems on soil organic carbon content and the density, biomass and community composition of earthworms after a ten year period. *Eur. J. Soil Biol.* 45 (3), 247–251.
- Falco, L.B., Sandler, R., Momo, R., Di Ciocco, C., Saravia, L., Coviella, C., 2015. Earthworm assemblages in different intensity of agricultural uses and their relation to edaphic variables. *PeerJ Preprints* 3, 1–18.
- Fusaro, S., Gavinelli, F., Lazzarini, F., Paoletti, M.G., 2018. Soil biological quality index based on earthworms (QBS-e). A new way to use earthworms as bioindicators in agroecosystems. *Ecol. Indic.* 93, 1276–1292.
- García-Pérez, J.A., Alarcón-Gutiérrez, E., Perroni, Y., Barois, I., 2014. Earthworm communities and soil properties in shaded coffee plantations with and without application of glyphosate. *Appl. Soil Ecol.* 83, 230–237.
- Gee, G.W., Bauder, J.W., 1986. Particle-size analysis. p. 383–411. Klute, A. (ed.), *Methods of soil analysis. Part 1. Physical and Mineralogical Methods*. Agronomy Monograph N° 9, second ed. ASA and SSSA, Wisconsin, USA.
- Godoy, J.L., Vega, J.R., Marchetti, J.L., 2014. Relationships between PCA and PLS-regression. *Chemom. Intell. Lab. Syst.* 130, 182–191.
- Jackson, M.L., 1976. Determinación de los cationes metálicos canjeables de los suelos. *Análisis químico de suelos*, ed. Omega S.A., Barcelona, España, p. 662.
- Jiménez, J.J., Lavelle, P., Decaens, T., 2006. The efficiency of soil hand-sorting in assessing the abundance and biomass of earthworm communities. Its usefulness in population dynamics and cohort analysis studies. *Eur. J. Soil Biol.* 42, 225–230.
- Jiménez, J.J., Decaens, T., Lavelle, P., Rossi, J.P., 2012. Soil environmental heterogeneity allows spatial co-occurrence of competitor earthworm species in a gallery forest of the Colombian “Llanos” *Oikos* 121 pp. 915–926.
- Jones, C.G., Lawton, J.H., Shachak, M., 1994. Organisms as ecosystem engineers. *Oikos* 69, 373–386.
- Johnston, A.S., Sibly, R.M., Hodson, M.E., Alvarez, T., Thorbek, P., 2015. Effects of agricultural management practices on earthworm populations and crop yield: validation and application of a mechanistic modelling approach. *J. Appl. Ecol.* 52 (5), 1334–1342.
- Johnston, A.S., Sibly, R.M., Thorbek, P., 2018. Forecasting tillage and soil warming effects on earthworm populations. *J. Appl. Ecol.* 55 (3), 1498–1509.
- Kanianska, R., Jad'ud'ová, J., Makovníková, J., Kizeková, M., 2016. Assessment of

- relationships between earthworms and soil abiotic and biotic factors as a tool in sustainable agricultural. Sustainability 8, 1–14.
- Kherbouche, D., Bernhard-Reversat, F., Moali, A., Lavelle, P., 2012. The effect of crops and farming practices on earthworm communities in Soummam valley. Algeria. Eur. J. Soil Biol. 48, 17–23.
- Lavelle, P., 1981. Stratégies de reproduction chez les vers de terre. Acta Oecol. 2, 117–133.
- Lavelle, P., Caro, G., Hartmann, C., Decaëns, T., Barot, S., Mora, P., Mathieu, J., 2007. Earthworms as key actors in self-organized soil systems. In: Cuddington, K. (Ed.), Ecosystem Engineers. Academic Press, pp. 77–107.
- Lemtiri, A., Colinet, G., Alabi, T., Cluzeau, D., Zirbes, L., Haubruge, E., Francis, F., 2014. Impacts of earthworms on soil components and dynamics. A review. Biotechnol. Agron. Soc. Environ. 18 (1), 121–133.
- Lewis, J.P., Collantes, M.B., 1974. La vegetación de la provincia de Santa Fe. Reseña general y enfoque del problema. Boletín de la Sociedad Argentina de Botánica 15, 343–356. <https://botanicaargentina.com.ar/wp-content/uploads/2018/09/343-356004.pdf>.
- MAG (Ministerio de Agricultura y Ganadería), 1982. Toma de muestras y determinaciones analíticas en suelos y aguas. Dirección General de Extensión e Investigación Agropecuaria. Departamento de Apoyo Analítico. Provincia de Santa Fe. Argentina.
- Maitre, M.I., Rodríguez, A.R., Masin, C.E., Ricardo, T., 2012. In: Pesticides-Advances in Chemical and Botanical Pesticide, pp. 382.
- Martens, H., Naes, T., 1989. Multivariate Calibration. John Wiley & Sons.
- Masin, C.E., Rodríguez, A.R., Maitre, M.I., 2011. Evaluación de la abundancia y diversidad de lombrices de tierra en relación con el uso del suelo en el Cinturón Hortícola de Santa Fe (Argentina). Ciencia del Suelo 29 (1), 21–28.
- Masin, C.E., 2017. Efectos de largo plazo del uso del suelo sobre la comunidad de lombrices de tierra (Annelida, Oligochaeta) en la provincia de Santa Fe. Doctoral thesis: UNL, Santa Fe, Argentina. <http://bibliotecavirtual.unl.edu.ar/tesis/handle/11185/937>.
- Miretti, M.C., Pilatti, M., Lavado, R.S., Imhoff, S.C., 2012. Historia de uso del suelo y contenido de micronutrientes en Argiudoles del centro de la provincia de Santa Fe (Argentina). Ciencia del Suelo 30 (1), 67–73.
- Mischis C.C. 1991., Las lombrices de tierra (Annelida, Oligochaeta) de la provincia de Córdoba, Argentina. Boletín de la Academia Nacional de Ciencias de Córdoba 59 (3 y 4), 187–237. <https://catalogo.biblio.unc.edu.ar/Record/agropecuarias.2643>.
- Ortiz-Gamino, D., Pérez-Rodríguez, P., Ortiz-Ceballos, A., 2016. Invasion of the tropical earthworm *Pontoscolex corethrurus* (Rhinodrilidae, Oligochaeta) in temperate grasslands. PeerJ 1–20.
- Paoletti, M.G., Sommaggio, D., Fusaro, S., 2013. Proposta di indice di qualità biologica del suolo (QBS-e) basato sui lombrichi e applicato agli agroecosistemi. Biol. Ambientale 27 (2), 25–43.
- Parisi, V., 2001. La qualità biologica del suolo. Un metodo basato sui microartropodi. Acta Naturalia de l'Ateneo Parmense 37, 87–106.
- Parisi, V., Menta, C., Gardi, C., Jacomini, C., Mozzanica, E., 2005. Microarthropod communities as a tool to assess soil quality and biodiversity: a new approach in Italy. Agric. Ecosyst. Environ. 105, 323–333.
- Pelosi, C., Bertrand, M., Roger-Estrade, J., 2009. Earthworm community in conventional, organic and direct seeding with living mulch cropping systems. Agron Sustain Dev. 29, 287–295.
- Pèrès, G., Vandenbulcke, F., Guernion, M., Hedde, M., Beguiristain, T., Douay, F., Houot, S., Piron, D., Richard, A., Bispo, A., Grand, C., Galsomies, L., Cluzeau, D., 2011. Earthworm indicators as tools for soil monitoring, characterization and risk assessment. An example from the national Bioindicator programme (France). Pedobiologia 54, 77–87.
- Reynolds J.W., 1996. Earthworm biology and ecology. Course Manual. Lindsay: Sir Sandford Fleming College, pp. 196.
- Roarty, S., Hackett, R.A., Schmidt, O., 2017. Earthworm populations in twelve cover crop and weed management combinations. Appl. Soil Ecol. 14, 142–151.
- Rudisser, J., Tasser, E., Peham, T., Meyer, E., Tappeiner, U., 2015. The dark side of biodiversity: spatial application of the biological soil quality indicators (BSQ). Ecol. Ind. 53, 240–246.
- Sarandón, S.J., 1998. The development and use of sustainability indicators: a need for organic agriculture evaluation. XII International Scientific Conference IFOAM. Mar del Plata, Argentina. 135pp.
- Southwood, T.R.E., Henderson, P.A., 2000. Ecological Methods, third ed. Blackwell Science, Oxford, UK, pp. 575.
- Walkley, A., Black, I.A., 1934. An examination of the different method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Sci. 37 (1), 29–38.
- Xie, T., Wang, M., Chen, W., Uwizeyimana, H., 2018. Impacts of urbanization and landscape patterns on the earthworm communities in residential areas in Beijing. Sci. Total Environ. 626, 1261–1269.