Hypotheses presence and acceptance in seven soil science journals

Pierfrancesco Nardi a,⁎, Giovanni Di Matteo a, Alfred E. Hartemink b

a Consiglio per la Ricerca e la sperimentazione in Agricoltura - Agricultural Research Council, Research Unit for Climatology and Meteorology applied to Agriculture, Via del Caravita, 7/a 00186 Rome, Italy
b University of Wisconsin — Madison, Department of Soil Science, FD Hole Soils Lab, 1525 Observatory Drive, Madison, WI 53706, USA

A R T I C L E   I N F O

Article history:
Received 14 May 2014
Received in revised form 5 December 2014
Accepted 7 December 2014
Available online xxxx

Keywords:
Hypotheses
Soil science
Soil journals
Hypotheses acceptance

A B S T R A C T

The presence and acceptance rates of hypotheses of papers in seven major soil journals were analyzed between 2001 and 2013. The aim of the study was to quantify the testing of hypotheses in soil science and investigate how it evolved over time. The journals were Applied Soil Ecology, Biology and Fertility of Soils, European Journal of Soil Science, Geoderma, Plant and Soil, Soil Biology and Biochemistry and Soil & Tillage Research. In total 15,344 papers were published by the seven journals over that period. Of a sample of 620 papers, 74% tested one hypothesis, 20% tested two or more hypotheses and 6% proposed a hypothesis. In total 66% of the all tested hypotheses (n = 783) were accepted, and the acceptance rate for the seven journals was more or less constant over time. A single hypothesis is more likely to be accepted (75%) compared to research with multiple hypotheses (55%). Although there was some difference between journals, it was concluded that acceptance rates of hypotheses in soil science are relatively low compared to other scientific disciplines.

C O R R E S P O N D I N G   A U T H O R

⁎ Corresponding author.
E-mail addresses: pierfrancesco.nardi@entecra.it (P. Nardi), hartemink@wisc.edu (A.E. Hartemink).

1. Introduction

The term hypothesis has different meanings in science. A hypothesis is viewed as the antecedent of the if–then statement or as a speculation (Guthery et al., 2004), as an explanation of an observe pattern (Krebs, 2000) for an observation, phenomenon, or scientific problem that can be tested by further investigation (Wolff and Krebs, 2008), or as an imaginative conjecture representing the first stage of scientific enquiry (Ayala, 2009). According to Earman and Salmon (1992), a hypothesis is a statement that is intended for evaluation in terms of its consequences. Garton et al. (2005) viewed a hypothesis as a prediction or deduction from a given theory. Ford (2000) defined hypothesis from postulate (an unexplored or a new idea) which is a data statement constructed to give a logical test. Hypotheses are classified as experimental, if they can be tested by field or laboratory experiments, or historical if they refer to past causes, unlikely to be reproduce through an experiment for currently observable phenomena. Historical hypotheses are recognized in geology (Cleland, 2001, 2002), biology (Simpson, 1967), and in soil science (Phillips, 2000). The differences in definition have led to incorrect use and a hypothesis is sometimes used as synonym of theory or theory’s axioms or as synonym for postulation. In this paper, we will refer to a hypothesis as a general idea that needs to be tested and that can be confirmed or rejected.

The testing of hypotheses according to the hypothetic-deductivism scheme is considered to be one of the standard methods of reasoning in science (Ayala, 2009; Murray, 2001). According to Popper’s (1959) falsificationist approach, theories can only be disproved and not proved and new knowledge arises by eliminating false theory and tests should be selected based on their capability to yield a disconfirming result instead confirming ones. The view of science as theory driven, as well as the use of hypothetic-deductive method of reasoning based on the falsification approach has been fairly well established in soil science. For example, J. von Liebig (cit in Azzone, 1991) the German soil chemist, affirmed that the natural science method is deductive and a priori which means that any experiment should be supported by a theory and therefore an experiment is significant only if it tests a theory. Prosser et al. (2007), highlighted the importance of theory for soil microbial ecology. Andrén et al. (2008) stated that the complexity and diversity of soil systems make it difficult to test hypotheses in a rigorous way and by affirming that there must be a way to reject a hypothesis, the authors implicitly invoked the Popper’s principle of falsification. Phillips (2008) discussed how models can generate hypothesis that can be tested by field-observation and he identified testable hypothesis with falsifiable hypothesis. Bradford and Fierer (2012) attempted to apply a falsification approach toward hypotheses that are relevant in the biogeography of microbial communities (e.g. functional redundancy, similarity and equivalence hypotheses). However, the falsification approach has been challenged by Thomas Kuhn (1970), who introduced the concepts of paradigm and revolution in science and pointed out how scientists usually do not process hypotheses according to the falsification principle.

The prevalence of papers supporting the tested hypotheses among scientific disciplines has been fairly well studied (Fanelli, 2012) and the higher acceptance rate of the tested hypotheses is considered too
good to be true (Ioannidis, 2005). Studies with significant results were more likely to be published in journals with a high impact factor (Easterbrook et al., 1991). On the other hand, papers reporting negative findings (the tested hypothesis was not supported by evidence) are more likely published in journals with lower impact factor (Littner et al., 2005). In addition, the tendency for researchers to submit manuscripts and of editors to accept papers based on the strength and outcome of the research findings is well known (Sterling, 1959; Sterling et al., 1995; Chalmers, 1990). Therefore, the ideal of science as objective as possible, theory-driven and falsificationist in spirit, may be in contrast to the reality of science that is biased in reporting experimental results.

The number of soil science publications is steadily increasing and so is the impact of most journals (Hartemink, 2001; Minasny et al., 2010). In part this reflects the vibrance of the discipline (Hartemink and McBratney, 2008), in part it is due to the "publish or perish" culture that seems to influence universities and research centers across the world. The rise in soil science publications has been fairly well quantified but the analysis of the testing of hypothesis remains to be investigated.

Here, we have researched the testing of hypotheses in soil science papers to address the following questions: (1) is the testing of hypothesis dominated by confirmation?, (2) are there differences in terms of hypothesis testing outcome between soil science and other scientific disciplines?, (3) can testing one or multiple hypotheses affects the hypothesis testing outcome?, and (4) how the testing of scientific hypotheses evolved with time?

This research quantified the testing of hypotheses in soil science and investigates how it evolved over the time. A survey was conducted for seven major soil journals (Applied Soil Ecology, Biology and Fertility of Soils, European Journal of Soil Science, Geoderma, Plant and Soil, Soil Biology and Biochemistry, Soil & Tillage Research). In total 655 papers were reviewed over the period 2001–2013. A systematic analysis of soil hypotheses in soil science has not been done before.

2. Data and analysis
2.1. Data collection

The survey covered a period of 13 years from 2001 to 2013 and was based on seven soil journals that are representative of different subdisciplines: Applied Soil Ecology, Biology and Fertility of Soils, European Journal of Soil Science, Geoderma, Plant and Soil, Soil Biology and Biochemistry and Soil & Tillage Research. The survey was conducted using the Elsevier’s Scopus database and was performed in two steps. Firstly, a total of 15,344 articles were identified using the journal ISSN (International Standard Serial Number) codes. We then identified hypothesis papers using the following search criteria: ISSN (journal) AND TITLE-ABS-KEY (hypothesis) OR hypotheses OR hypothesize OR hypothesise OR hypothesized OR hypothesised AND DOCTYPE (ar).

In total 969 hypothesis papers were found. To determine the sample size of papers to be analyzed we used the equation (Cochran, 1963):

\[ n_0 = \frac{Z^2 \times p \times q}{e^2} \]  

(1)\n
where \( n_0 \) = sample size, \( Z^2 \) is the abscissa of the normal curve that cuts off an area \( \alpha \) at the tails and was equal to 1.96. The \( e \) is the accepted sampling error and was equal to 0.05, whereas \( p \) is the estimated proportion of an attribute in the population (e.g. accepted and rejected hypotheses). As we had no a priori knowledge of these proportions among the seven journals, we fixed \( p = 0.5 \) assuming therefore the maximum variability.

However, as the seven journals and related papers corresponded to seven different finite populations, we used:

\[ n = \frac{n_0}{1 + \frac{n_0-1}{N}} \]  

(2)\n
for finite population correction adjustment where \( n \) is the adjusted sample size and \( N \) is the population size. That is: 83 for Applied Soil Ecology, 54 for Biology and Fertility of Soils, 51 for European Journal of Soil Science, 95 for Geoderma, 270 for Plant and Soil, 343 for Soil Biology and Biochemistry and 73 for Soil & Tillage Research. In total, 655 papers containing at least one hypothesis were analyzed.

By reading the abstract or the full text, they were classified as follows:

1. papers that tested one hypothesis (single hypothesis papers);
2. papers that tested two or more hypotheses (multiple hypotheses papers);
3. papers that proposed a hypothesis (hypothesis formulation papers).

In the first two types of papers, the tested hypotheses were classified as accepted or rejected. A hypothesis was considered accepted according to what authors declared in the abstract or discussion sections. For example, sentences like "our data support the hypothesis..." or "our hypothesis is accepted" or "we verify our hypothesis" or "our hypothesis is confirmed...". On the contrary, when authors clearly stated that they did not find any support for the stated hypothesis like "data do not support our first hypothesis" or "we rejected our hypothesis that..." or "our hypothesis cannot be confirmed" a hypothesis was considered rejected.

In case of single hypothesis papers, the number of papers corresponds to the number of hypothesis (one paper = one hypothesis). For papers with multiple hypotheses the number of hypotheses is higher than that of papers. To avoid bias that could alter proportions between hypotheses accepted and rejected, all tested hypotheses (single or multiple, \( n = 783 \)) were considered. From the 655 randomly selected papers, 35 were excluded because of the lack of clarity of the outcomes and all analyses were done on the remaining 620 papers.

We considered only research papers while short papers, reviews, letters and other types were excluded.

2.2. Data analysis

We have listed the number of hypotheses as counts and percentages. To test the association between dependent (acceptance and rejection of the tested hypothesis) and independent variables (e.g. single hypothesis or multiple hypotheses papers, journals) the chi-square test (\( \chi^2 \)) was used. In the case of n by n (n > 2) contingency tables when the chi-square was significant, an analysis of adjusted residuals was performed to determine which cells were the major contributors to the chi-square significance (Sheskin, 2000). By definition, a residual is the difference in the observed frequency and the expected frequency. Adjusted residuals are then calculated as follows: standardized residuals / estimated standard error. Adjusted residuals are approximately normally distributed with a mean of 0 and a standard deviation of 1. Adjusted residuals with an absolute value that is equal to or greater than the tabled critical two-tailed 0.05 (\( Z_{0.05} \sim 1.96 \)) or 0.01 (\( Z_{0.01} \sim 2.58 \)) values are significant at 0.05 and 0.01 levels respectively. The sign of the adjusted residual indicates whether the observed frequency of the cell is above (+) or below (−) the expected frequency (Sheskin, 2000).

In addition to the chi-square test, that is dependent of sample size, we incorporate a sample size independent measure like odds. For the two by two contingency tables, odds, odds ratios (OR) and a 95% confidence interval were calculated. The odds of something happening are how many times more likely the event happen are to something not happening. In particular, odds were calculated by dividing the frequency of
accepted to rejected hypothesis for single and multiple hypotheses papers separately. Odds ratios represent the ratio of two odds. The odds ratio expresses the degree of association between the two variables in a different numerical format than chi-square test and provides a complementary way to interpret the results of the contingency tables. For OR = 1, an event is equally likely under both situations. If OR > 1, the event with the first odds is more likely. If OR < 1, the event with the second odds is more likely (Sokal and Rohlf, 1995). Statistical significances of the OR were calculated using a 95% confidence interval. The determination of odds and odds ratios was limited to a 2 × 2 table due to the known difficulty to interpret them in larger contingency tables. All statistical analyses were carried out using SPSS version 19 software.

3. Results

3.1. Papers with hypotheses

A total of 15,344 soil science papers were published by the seven journals in the period 2001–2013. About one-quarter of all papers were published in *Plant and Soil* (n = 3882) or *Soil Biology and Biochemistry* (n = 3632). *Geoderma* published about 2600 papers while the remaining journals published papers ranging from 1043 for the *European Journal of Soil Science* to 1494 for *Soil & Tillage Research*. Of all published papers, only 6% (n = 969) tested or proposed at least one hypothesis.

For all the seven soil science journals, there were 3.6 times more papers that tested one hypothesis compared to papers that tested two or more hypotheses (multiple hypotheses papers) (Table 1). *Soil & Tillage Research* had the highest number of paper (87%) with single hypothesis papers whereas 27% of the papers published in *Biology and Fertility of Soils* had multiple hypotheses. Hypothesis formulation papers were in the range of 3% for *Soil Biology and Biochemistry* and *Soil & Tillage Research* to 9% for *Applied Soil Ecology*. The chi-square test showed a significant association between types of paper and journals ($\chi^2 = 24.861, df = 12, p < 0.05$) (Table 2). Residual analysis showed that single hypothesis papers were over-represented for *Soil & Tillage Research* (AR = 2.47). The frequency of multiple hypothesis papers was significantly lower for *Applied Soil Ecology* (AR = −2.02), *Geoderma* (AR = −2.13) and for *Soil & Tillage Research* (AR = −2.20) and significantly higher for *Soil Biology and Biochemistry* (AR = 2.93). No significant differences between journals emerged for hypothesis formulation papers.

3.2. Acceptance of hypotheses

Of the 783 tested hypotheses, 66% (n = 520) were accepted by the authors. In all the seven journals, proportions of accepted over rejected hypotheses were higher in papers that tested one single hypothesis than in multiple hypotheses papers, as well as in the total of single plus multiple hypotheses articles. The exception to this was *Soil & Tillage Research* that had a higher percentage (93%) of accepted hypotheses for multiple hypotheses articles (Table 3).

There was a significant association between the approach used to test the hypothesis (single vs multiple) and the testing hypothesis outcome ($\chi^2 = 32.614, df = 1, p < 0.0001$). However, when the test was performed separately for each of the seven journals, the association was not significant for *Applied Soil Ecology*, *European Journal of Soil Science* and *Soil & Tillage Research* and significant for the remaining four journals.

Odds, and OR with 95% confidence interval are shown in Table 3. For all the seven journals, a hypothesis was about three times (Odds = 2.923) as likely to be confirmed as rejected when tested singularly. To the contrary, the acceptance odds for multiply hypotheses were only 1.219. The derived OR, useful to test differences between testing hypothesis outcomes according to the type of papers (single and multiple hypotheses papers), showed that the odds that a hypothesis will be accepted are more than two times greater if tested singularly (OR = 2.397, 95% CI = 1.771–3.247). *Geoderma* was the journal with the highest and significant OR (OR = 6.308, 95% CI = 1.382–26.843), while *Soil & Tillage Research* showed the lowest but not significant OR (OR = 0.279, 95% CI = 0.033–2.359).

A chi-square test based on a 7 × 4 (7 = number of journals; 4 = possible responses that is accepted and rejected hypotheses from single

<table>
<thead>
<tr>
<th>Journal</th>
<th>Total number of papers</th>
<th>Number of papers analyzed</th>
<th>% of papers analyzed</th>
<th>Number of papers classified</th>
<th>% of papers classified</th>
<th>Single hypothesis (%)</th>
<th>Multiple hypotheses (%)</th>
<th>Hypothesis formulation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied Soil Ecology</td>
<td>83</td>
<td>69</td>
<td>83</td>
<td>65</td>
<td>94</td>
<td>80</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Biology and Fertility of Soils</td>
<td>54</td>
<td>52</td>
<td>96</td>
<td>48</td>
<td>92</td>
<td>65</td>
<td>27</td>
<td>8</td>
</tr>
<tr>
<td>European Journal of Soil Science</td>
<td>51</td>
<td>47</td>
<td>92</td>
<td>42</td>
<td>89</td>
<td>74</td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td>Geoderma</td>
<td>95</td>
<td>76</td>
<td>80</td>
<td>67</td>
<td>88</td>
<td>81</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Plant and Soil</td>
<td>270</td>
<td>163</td>
<td>60</td>
<td>160</td>
<td>98</td>
<td>73</td>
<td>22</td>
<td>6</td>
</tr>
<tr>
<td>Soil Biology and Biochemistry</td>
<td>343</td>
<td>184</td>
<td>54</td>
<td>176</td>
<td>96</td>
<td>69</td>
<td>28</td>
<td>3</td>
</tr>
<tr>
<td>Soil &amp; Tillage Research</td>
<td>73</td>
<td>64</td>
<td>88</td>
<td>62</td>
<td>97</td>
<td>87</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>969</td>
<td>655</td>
<td>68</td>
<td>620</td>
<td>95</td>
<td>74</td>
<td>20</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Journals</th>
<th>Single hypothesis paper</th>
<th>Multiple hypotheses paper</th>
<th>Hypothesis formulation paper</th>
<th>$\chi^2$ (df = 12)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Residuals</td>
<td>AR</td>
<td>Residuals</td>
<td>AR</td>
<td>Residuals</td>
</tr>
<tr>
<td>Applied Soil Ecology</td>
<td>3.88</td>
<td>1.16</td>
<td>-6.21</td>
<td>-2.02</td>
<td>2.33</td>
</tr>
<tr>
<td>Biology and Fertility of Soils</td>
<td>-4.54</td>
<td>-1.55</td>
<td>3.25</td>
<td>1.21</td>
<td>1.29</td>
</tr>
<tr>
<td>European Journal of Soil Science</td>
<td>-0.09</td>
<td>-0.03</td>
<td>0.46</td>
<td>0.18</td>
<td>-0.37</td>
</tr>
<tr>
<td>Geoderma</td>
<td>4.40</td>
<td>1.30</td>
<td>-6.62</td>
<td>-2.13</td>
<td>2.22</td>
</tr>
<tr>
<td>Plant and Soil</td>
<td>-2.45</td>
<td>-0.51</td>
<td>2.48</td>
<td>0.57</td>
<td>-0.03</td>
</tr>
<tr>
<td>Soil Biology and Biochemistry</td>
<td>-9.30</td>
<td>-1.89</td>
<td>13.23</td>
<td>2.93</td>
<td>-3.94</td>
</tr>
<tr>
<td>Soil &amp; Tillage Research</td>
<td>8.10</td>
<td>2.47</td>
<td>-6.60</td>
<td>-2.20</td>
<td>-1.50</td>
</tr>
</tbody>
</table>
and multiple hypotheses papers) showed a significant relationship between outcomes and journals ($\chi^2 = 62.334$, df = 18, $p < 0.0001$) (Table 4). The analysis of adjusted residuals for single hypothesis papers showed that accepted hypotheses were significantly over-represented for Geoderma (AR = 3.60) and Soil & Tillage Research (AR = 3.30) and under-represented for Soil Biology and Biochemistry (AR = −3.30) journals. In the case of multiple hypotheses papers, Geoderma (AR = −3.50) and Soil Biology and Biochemistry (AR = 2.00), had significant lower and higher frequency of acceptance of hypotheses. The rejection of multiple hypotheses was significantly higher for Biology and Fertility of Soils (AR = 2.00) and Soil Biology and Biochemistry (AR = 2.80) and lower in Soil Tillage & Research (AR = −3.80). No differences were found for the other journals.

### 3.3. Temporal trends

Trends over time in total number of papers with hypothesis and their proportions (with respect to the total number of published papers) are presented in Fig. 1. The number of papers with a hypothesis gradually decreased between 2001 ($n = 63$) and 2005 ($n = 46$). After 2005, the number of papers with a hypothesis increased and exceeded 100 articles in 2010 and 2013. The proportion of hypothesis papers remained more or less constant for the whole period and ranged between 5 and 8%. The share of hypothesis papers over the time is similar for most journals (Fig. 1) with the exception of Soil & Tillage Research that showed an increasing trend in the last four years and the share of papers with one or more hypothesis increased to 17% in 2010.

The share of accepted hypotheses remained more or less constant over time (Fig. 2). The chi-square test showed no significant association between the year in which the papers were published and the rejection of acceptance of hypothesis for both single and multiple hypotheses papers ($\chi^2 = 13.166$, df = 12, $p < 0.05$) or when single hypothesis, multiple hypotheses, and multiple hypotheses papers were analyzed separately ($\chi^2 = 13.166$, df = 12, $p < 0.05$) and multiple hypotheses papers were analyzed separately ($\chi^2 = 13.166$, df = 12, $p < 0.05$) or when single hypothesis, multiple hypotheses, and multiple hypotheses papers were analyzed separately ($\chi^2 = 13.166$, df = 12, $p < 0.05$) and multiple hypotheses papers were analyzed separately ($\chi^2 = 13.166$, df = 12, $p < 0.05$). The journal Applied Soil Ecology showed a gradual increase in hypotheses acceptance over the period 2005–2013 and some hypotheses rejection in 2004 and 2006. Such fluctuation between acceptance and rejection of hypotheses was also observed in Biology and Fertility of Soils. The European Journal of Soil Science hypotheses acceptance peaked to 100% in several years and that was observed two times for Geoderma. The hypotheses acceptance for Plant and Soil was in the range of 50% to 77% with a peak of 88% in 2006, whereas Soil Biology and Biochemistry was characterized by a lower acceptance (32%) in 2002.

![Figure 1: Temporal trends in hypothesis acceptance](image1)

**Table 3**

<table>
<thead>
<tr>
<th>Journal</th>
<th>% of accepted</th>
<th>% of rejected</th>
<th>Number of hypotheses</th>
<th>Odds</th>
<th>OR (95% CI)</th>
<th>$\chi^2$ (df = 1)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied Soil Ecology</td>
<td>Single</td>
<td>71</td>
<td>29</td>
<td>52</td>
<td>2.467</td>
<td>2.220 (0.752–6.551)</td>
<td>2.134</td>
</tr>
<tr>
<td></td>
<td>Multiple</td>
<td>53</td>
<td>47</td>
<td>19</td>
<td>1.111</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>66</td>
<td>64</td>
<td>71</td>
<td>1.958</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biology and Fertility of Soils</td>
<td>Single</td>
<td>77</td>
<td>23</td>
<td>31</td>
<td>3.429</td>
<td>4.114 (1.390–12.182)</td>
<td>6.861</td>
</tr>
<tr>
<td></td>
<td>Multiple</td>
<td>45</td>
<td>55</td>
<td>33</td>
<td>0.833</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>61</td>
<td>98</td>
<td>64</td>
<td>1.560</td>
<td></td>
<td></td>
</tr>
<tr>
<td>European Journal of Soil Science</td>
<td>Single</td>
<td>84</td>
<td>16</td>
<td>31</td>
<td>5.200</td>
<td>3.343 (0.937–11.924)</td>
<td>3.638</td>
</tr>
<tr>
<td></td>
<td>Multiple</td>
<td>61</td>
<td>39</td>
<td>23</td>
<td>1.556</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>74</td>
<td>55</td>
<td>54</td>
<td>2.857</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil Biology and Biochemistry</td>
<td>Single</td>
<td>70</td>
<td>30</td>
<td>116</td>
<td>2.515</td>
<td>1.956 (1.106–3.460)</td>
<td>5.38</td>
</tr>
<tr>
<td></td>
<td>Multiple</td>
<td>96</td>
<td>35</td>
<td>96</td>
<td>1.286</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>166</td>
<td>65</td>
<td>212</td>
<td>1.827</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil &amp; Tillage Research</td>
<td>Single</td>
<td>72</td>
<td>28</td>
<td>116</td>
<td>2.515</td>
<td>1.956 (1.106–3.460)</td>
<td>5.38</td>
</tr>
<tr>
<td></td>
<td>Multiple</td>
<td>96</td>
<td>44</td>
<td>96</td>
<td>1.286</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>168</td>
<td>116</td>
<td>212</td>
<td>1.827</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All seven journals</td>
<td>Single</td>
<td>75</td>
<td>25</td>
<td>121</td>
<td>2.667</td>
<td>2.393 (1.412–4.059)</td>
<td>10.659</td>
</tr>
<tr>
<td></td>
<td>Multiple</td>
<td>55</td>
<td>45</td>
<td>129</td>
<td>1.115</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>66</td>
<td>70</td>
<td>250</td>
<td>1.660</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\chi^2 =$ chi-square; df = degree of freedom; OR (95% CI) = odds ratios (95% confidence interval).

a 'Yates' correction for continuity.

## Table 4

Analysis of residuals for seven major soil science journals (single vs multiple hypotheses papers). Numbers in bold indicate significant ($p < 0.05$) adjusted residuals. AR = adjusted residuals.

<table>
<thead>
<tr>
<th>Journals</th>
<th>Single hypotheses</th>
<th>Multiple hypotheses</th>
<th>$\chi^2$ (df = 18)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accepted Residuals</td>
<td>Rejected Residuals</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Residuals AR</td>
<td>Residuals AR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applied Soil Ecology</td>
<td>6.00</td>
<td>1.50</td>
<td>−6.10</td>
<td>−1.80</td>
</tr>
<tr>
<td></td>
<td>4.40</td>
<td>1.50</td>
<td>−4.20</td>
<td>−1.40</td>
</tr>
<tr>
<td></td>
<td>1.50</td>
<td></td>
<td>62.334</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>Biology and Fertility of Soils</td>
<td>−3.95</td>
<td>−1.04</td>
<td>0.50</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>−2.60</td>
<td>−0.90</td>
<td>6.10</td>
<td>2.00</td>
</tr>
<tr>
<td>European Journal of Soil Science</td>
<td>2.40</td>
<td>0.70</td>
<td>1.70</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>−3.10</td>
<td>−1.20</td>
<td>−1.10</td>
<td>0.40</td>
</tr>
<tr>
<td>Geoderma</td>
<td>13.50</td>
<td>3.60</td>
<td>11.30</td>
<td>−3.50</td>
</tr>
<tr>
<td></td>
<td>3.60</td>
<td>1.30</td>
<td>−5.50</td>
<td>−1.90</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant and Soil</td>
<td>−9.60</td>
<td>−1.60</td>
<td>5.80</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1.30</td>
<td>0.30</td>
<td>2.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Soil Biology and Biochemistry</td>
<td>−21.20</td>
<td>−3.30</td>
<td>11.20</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>−4.40</td>
<td>−0.90</td>
<td>14.40</td>
<td>2.80</td>
</tr>
<tr>
<td>Soil &amp; Tillage Research</td>
<td>12.90</td>
<td>3.30</td>
<td>−1.70</td>
<td>−0.50</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>0.20</td>
<td>−11.90</td>
<td>−3.80</td>
</tr>
</tbody>
</table>

$\chi^2 =$ chi-square; df = degree of freedom; OR (95% CI) = odds ratios (95% confidence interval).
4. Discussion

To our knowledge, this survey of seven major soil science journals is the first study quantifying and analyzing the testing of hypotheses in soil science. We have counted the number of papers with one or more hypotheses and investigated trends over time. Differences were found between the journals and over time. In this discussion, we shall focus on a possible data explanation of some of the differences and similarities, as well as compare our results to similar studies in other scientific disciplines.

4.1. Hypothesis papers

It was found that only 6% of the 15,344 soil science papers published between 2001 and 2013, tested or formulated at least one hypothesis. This finding is in agreement with Desbiens (2006), who found that in the two prestigious journals *Nature* and *Science*, only 10% of papers published for the year 2004 explicitly stated a hypothesis. An informal survey conducted in wildlife research literature (only one journal analyzed), the percentages of hypothesis papers were 0% in 1971, 2.5% in 1981 and about 25% in 1991 and 2001 (Guthery et al., 2004).

However the low frequency of hypothesis papers in soil science led us to the question why so many soil papers lack hypotheses. We can think of three possible explanations.
Firstly, those papers that stated a hypothesis at the end of the introduction but not in the title/abstract/keywords sections were not retrieved in our survey because there is no way to search for a specific keyword (e.g. hypothesis) in the introduction section of the Scopus database. This is a limit of our survey and the accuracy of the estimated number of hypothesis papers could be improved by a future literature analysis. However, authors tend to report the most important research findings in the title or in the abstract. As we used keywords in a systematic way similar to other studies (Kyzas et al., 2007; Pautasso, 2010; de Winter and Dodou, 2014; Fanelli, 2012), it allows for the comparisons between soil science and similar studies in other disciplines.

Secondly, we found that soil scientists tend to use different wording for the term hypothesis (e.g. theory, model, prediction) and those are excluded from the survey. However, through an additional analysis...
we conducted on the same seven soil journals, using the same search criteria but using the words “theory”", “postulate” AND NOT “hypothesis", we found an additional 300 papers (data not shown) that lead to a percentage of 8%. 

Lastly, soil scientists may leave research questions and assumptions without translating them into propositions (e.g. hypothesis) that can be judged as true or false, leading papers to be question or aims oriented. In addition, scientific questions are of “what, when and where” types that constitute observational or descriptive rather than test specific hypotheses (Loehle, 1987).

Although, it is possible to improve the number of hypothesis papers by screening all soil science papers over the same period (n = 15,344), adding more soil journals or by using other searching criteria (words) or databases, we think it would not largely change the outcome of this research.

Soil science is a relatively young science (Brevik and Hartemink, 2010) and there used to be a slight prevalence of descriptive studies over theoretical ones. Young scientific disciplines move toward theory formulation rather than theory testing. A vast body of theories now exists for all soil disciplines and the apparent low number of hypothesis papers is likely due to cultural reasons. Soil science books are written in a manner that does not explicitly report the structure of theories. Soil processes, for example, are usually described in terms of “who” is responsible for or under what condition they vary, but an organization of all these information in terms of axioms, postulate, and data statements, that constitute the body of a theory is lacking. In other words, a theory is usually reported without a clear definition of its structure. We think that the knowledge of theory as well as its structure is preparatory to its application. The need for more theory-driven research is not new in soil science as it has already been highlighted for microbial ecology (Prosser et al., 2007).

The second part of this is that fundamental research in soil science has diminished and more and more applied research is being conducted (Hartemink, 2006). Many research projects are short-term (<3 years) and do not allow for rigid hypothesis testing based on sound theory.

4.2. Acceptance of hypotheses

About 66% of all tested hypotheses were accepted. There was some difference between journals and the acceptance rates of hypotheses were higher for Soil & Tillage Research and lower for Biology and Fertility of Soils and Soil Biology and Biochemistry journals. There is evidence that papers reporting positive or confirming results are more likely to be published and accepted by journals (Murtaugh, 2002) which may suggest a bias (Begg and Berlin, 1988). Papers reporting negative findings (rejection of hypothesis) are more difficult to publish (Dwan et al., 2008; Hasenboehler et al., 2007). Therefore, under scientific competition and the publishing pressure, scientists may wish to confirm their hypothesis and ignore contradictory evidences (Nickerson, 1998). The directionality of submitting manuscripts according to their outcomes distorts scientific literature, might dissuade high risk project and encourage scientists to fabricate and falsify their data (Fanelli, 2009).

We cannot exclude that there may be a slight bias in soil science papers as hypotheses are tested with a priori high probability to be confirmed. Although hypothesis papers are written according to the hypothetico-deductive scheme, that is a researcher first states a hypothesis and then tests it empirically through an experiment, a more acceptable hypothesis could be injected after the results are obtained. Kerr (1998), defined this questionable behavior as hypothesizing after the results are known (HARKing). In particular, HARKing is defined as presenting a post hoc hypothesis as if it was an a priori hypothesis. We think that HARKing may occur in soil science to some extent and consequently, some papers presented as hypothesis-driven could indeed be the results of exploratory studies.

However, soil science has the lowest percentage of acceptance among scientific disciplines. In medical sciences the acceptance of hypotheses has been found to be in the range of 72 to 74% (Dickersin et al., 1987; Dickersin, 1990; De Oliveira et al., 2012), whereas in psychology more than 95% of papers provided support for at least one of the tested hypotheses (Spence and Blanchard, 2001; Vasilev, 2013). This was also found for other scientific disciplines like space science with an acceptance rate of 70%, whereas Geosciences, Ecology, Agricultural Sciences, and Plant and Animal Sciences showed a percentage of accepted hypothesis ranging from about 75% to 85% (Fanelli, 2010b).

The method of multiple working hypotheses (Chamberlin, 1890), that represent one version of the hypothetico-deductive approach, has been applied to soil genesis by Arnold (1965) and Dijkerman (1974). Yet, using common garden and reciprocal transplant approaches, Strickland et al. (2009) tested the two competing hypotheses of functional equivalence and dissimilarity for litter mineralization. Our results suggest how the outcome of hypothesis testing is related to the number of hypotheses tested in the experiment. In fact, it was found that testing a single hypothesis is more likely to be confirmed compared to the testing of multiple hypotheses. This was particularly evident for journals like Biology and Fertility of Soils and Soil Biology and Biochemistry. An exception to this finding was Soil & Tillage Research, for which the frequency of accepted hypothesis decreased when multiple hypotheses were tested.

Hypothesis rejection is not synonymous for poor results or inadequate data. Therefore, an increasing number of studies embracing the multiple hypotheses approach could improve the rigor and objectivity and avoid potential questionable behaviors of testing research hypotheses in soil science.

4.3. Trends

Fanelli (2010a, 2012), found that hypotheses rejection rates are declining from most of 20 surveyed disciplines and that the proportion of papers reporting hypotheses acceptance has increased by more than 20% between 1990 and 2007. Pautasso (2010), reported that the ratio of positive to negative results has been increasing over time. Some disciplines like Geosciences and Plant and Animal Sciences for example, showed a steady or a slightly declining trend of hypotheses acceptance whereas disciplines like Agricultural Sciences, Microbiology, Molecular Biology, Biology and Biochemistry, showed a positive and sometimes strong increase of hypotheses acceptance (Fanelli, 2012).

We did not detect an increasing or decreasing trend. Our results seem to be in agreement to those of Fanelli (2012) for Geosciences and Plant and Animal Sciences.

The cost of a soil science publication has been estimated to be about €200,000 (Hartemink and McBramey, 2008), so reducing bias in scientific publications is necessary. Journal editors and soil scientists should play an active role to avoid potential bias. For example, editors should declare more explicitly how the outcome of the tested hypothesis is irrelevant to the acceptability of the submitted manuscript. Authors could highlight negative findings and give details why such outcomes are important.

5. Conclusions

From this research the following can be concluded:

- Less than 6% of the total soil published papers have a hypothesis.
- Soil science papers have a relatively low acceptance rate of tested hypothesis compared to other scientific disciplines and no temporal trends have been detected for the surveyed seven soil journals.
- Soil hypothesis outcome seems to be driven by the number of hypotheses tested by soil scientists. When multiple hypotheses were used, the rate of confirmed hypothesis decreases. When a hypothesis is tested singularly, it is more likely to be confirmed and this was found for all seven journals.
Acknowledgments

We thank three anonymous reviewers and the Editor-in-Chief for their valuable comments on the draft of this paper. Pierfrancesco Nardi receives funding from ERANET FORESTERRA project “Enhancing Forest RESearch in the MediTERRAnean through an improved coordination and integration” (grant number: 291832). We much enjoyed the research and had hypothesized that this paper would take far less time. It didn’t.

References


Fanelli, D., 2012. Negative results are disappearing from most disciplines and countries. Scientometrics 90, 891–904.


