

Artigo

Avaliação da qualidade do solo sob diferentes manejos Estratégias: Uso combinado de abordagens estatísticas para selecionar os indicadores físico-químicos do solo mais informativos

Anna Maria Stellacci ¹, Mirko Castellini ^{2,*} , Mariangela Deacon ² , Roberta Rossi ³
e Concetta Eliana Gattullo ¹ 

¹ Departamento de Ciências do Solo, Plantas e Alimentos, Universidade de Bari "A. Moro", 70126 Bari, Itália; annamaria.stellacci@uniba.it (AMS); concettaeliana.gattullo@uniba.it (CEG)

² Conselho de Pesquisa e Economia Agrícola - Pesquisa em Agricultura e Meio Ambiente, Centro (CREA-AA), 70125 Bari, Itália; mariangela.diacono@crea.gov.it

³ Conselho de Pesquisa Agrícola e Análise de Economia Agrícola - Centro de Pesquisa em Produção Animal e Aquicultura, Bella Muro, PZ, SS 7, 85051 Potenza, Itália; roberta.rossi@crea.gov.it

* Correspondência: mirko.castellini@crea.gov.it

Resumo: A avaliação da qualidade do solo sob diferentes práticas de manejo é crucial para a produção agrícola sustentável e o uso de recursos naturais. Neste estudo, diferentes métodos estatísticos (análise de componentes principais, PCA; análise discriminante passo a passo, SDA; regressão por mínimos quadrados parciais com estatística VIP, PLSR) foram aplicados para identificar as variáveis que mais discriminaram o estado do solo sob cultivo mínimo e plantio direto. Dados coletados em 2015 de um experimento de campo de longo prazo com trigo duro (*Triticum durum* Desf.) foram utilizados e vinte indicadores de solo (químicos, físicos e biológicos) foram quantificados para a camada superior do solo (0–0,20 m). A iteração de longo prazo de diferentes estratégias de manejo afetou a qualidade do solo, mostrando maior densidade do solo, capacidade relativa de campo (RFC), teores de carbono orgânico e extraível (COT e TEC) e potássio trocável sob plantio direto. PCA e SDA confirmaram esses resultados e sublinharam também o papel das frações de fósforo disponível e carbono orgânico como variáveis que mais discriminaram os tratamentos investigados. O PLSR, incluindo informações sobre a resposta da planta (rendimento de grãos e teor de proteína), selecionou, como as variáveis mais importantes, nutrientes da planta, indicadores de qualidade física do solo, pH e cátions trocáveis. A pesquisa mostrou a eficácia da combinação de métodos de seleção de variáveis para resumir informações derivadas de conjuntos de dados multivariados e melhorar a compreensão do sistema investigado. As abordagens estatísticas comparadas forneceram resultados diferentes em termos de variáveis selecionadas e a classificação das variáveis selecionadas. O uso combinado dos três métodos permitiu a seleção de um número menor de variáveis (COT, TEC, Olsen P, nitrogênio extraível em água, RFC, macroporosidade, capacidade de ar), que foram capazes de fornecer uma discriminação clara entre os tratamentos comparados, como mostrado pela ACP realizada no conjunto de dados reduzido. A presença de uma variável de resposta no PLSR impulsionou consideravelmente o processo de



Citação: Stellacci, AM; Castellini, M.; Diacono, M.; Rossi, R.; Gattullo, A. Avaliação da qualidade do solo por EC sob diferentes estratégias de manejo do solo: Uso combinado de estatísticas. *Abordagens para Selecionar os Indicadores Físico-Químicos do Solo Mais Informativos*. Appl. Sci. **2021**, *11*, 5099. <https://doi.org/10.3390/app11115099>

Editor acadêmico: Gerhard Soja

Recebido: 22 de abril de 2021

Aceito: 26 de maio de 2021

Publicado: 31 de maio de 2021

Nota do editor: O MDPI permanece neutro em relação a reivindicações jurisdicionais em mapas publicados e afiliações institucionais. iações.

Palavras-chave: seleção de variáveis; análise de componentes principais; análise discriminante passo a passo; regressão de mínimos quadrados parciais; estatística PLS-VIP; cultivo mínimo; plantio direto; experimentos de campo de longo prazo



Direitos autorais: © 2021 dos autores.

Licenciado MDPI, Basileia, Suíça.

Este artigo é um artigo de acesso aberto distribuído sob os termos e

condições do Creative Commons

Licença de atribuição (CC BY) (<https://creativecommons.org/licenses/by/4.0/>).

1. Introdução

O solo é cada vez mais reconhecido como um importante provedor de serviços ecossistêmicos [1,2]. A produção de alimentos e biomassa, a mitigação das mudanças climáticas por meio do ciclo e sequestro de carbono, a prevenção da degradação do solo, a purificação e o abastecimento de água são alguns dos serviços apoiados, regulados e fornecidos pelos solos [3,4]. Além disso, a qualidade do solo tem fortes implicações na saúde humana, tanto pela produção de alimentos seguros e nutritivos quanto pela proteção contra a poluição ambiental, demonstrando assim seu importante papel tanto na sociedade quanto no meio ambiente [5,6]. Essas evidências aumentaram a conscientização de

os recursos do solo devem ser protegidos da degradação e aumenta a necessidade de informações eficazes e confiáveis sobre modificações/alterações na qualidade do solo como consequência de seu manejo [4].

A atividade agrícola intensiva e as práticas de manejo não sustentáveis podem dar origem a processos de degradação do solo com consequente declínio da produtividade agrícola e da qualidade do solo [4,7] e uma redução associada nas funções e serviços ecossistêmicos [8,9]. Entre as práticas de uso da terra e manejo do solo, a lavoura afeta fortemente a qualidade física do solo, uma vez que modifica diretamente a agregação e a estabilidade estrutural do solo [10], a porosidade do solo [11,12] e as propriedades hidrológicas [13] e, portanto, a capacidade do solo de armazenar e transmitir água e ar [14,15]. Além disso, o manejo da lavoura desempenha um papel fundamental na renovação da matéria orgânica do solo, na abundância e diversidade microbiana, no armazenamento de carbono e, portanto, nas emissões de CO₂ [16,17]. Várias pesquisas descreveram os efeitos benéficos do manejo do plantio direto nas propriedades do solo [17–20]. No entanto, efeitos significativos nas mudanças de carbono do solo e na qualidade geral do solo são frequentemente difíceis de detectar no curto e médio prazo, devido às respostas em dois estágios à mudança exibidas por muitas propriedades fundamentais do solo [17,21]. Neste contexto, experimentos de campo de longo prazo podem ser considerados ferramentas de pesquisa ideais onde é provável que condições relativamente estáveis sejam observadas [12,15,21–23].

A qualidade do solo pode ser inferida pela identificação e medição dos indicadores de qualidade do solo, que são propriedades e processos específicos do solo sensíveis ao uso e manejo da terra [5,24–27]. A maioria dos autores utilizou indicadores únicos para avaliar a qualidade do solo e sua relação com os usos da terra [5]; no entanto, abordagens univariadas nem sempre permitem um julgamento abrangente sobre o estado do solo. Além disso, o aumento do número de indicadores pode aumentar a colinearidade ou fornecer resultados conflitantes, dificultando a avaliação da qualidade do solo [15,28]. A seleção e combinação de indicadores de diferentes naturezas (físicas, químicas e biológicas), por meio do cálculo de índices de qualidade do solo (IQS), é essencial para obter uma “imagem holística da qualidade do solo” [28]. Geralmente, a avaliação de um IQS compreende três etapas principais: (1) medição e seleção dos atributos mais relevantes do solo; (2) normalização e pontuação das propriedades selecionadas do solo usando funções de pontuação apropriadas; (3) integração dos atributos pontuados para construir o índice final [6,29,30]. A seleção das propriedades do solo mais relevantes é uma etapa crítica, pois constitui a base do índice de qualidade do solo. Os indicadores selecionados representam as variáveis mínimas do conjunto de dados [28]. Muitos estudos têm se concentrado na seleção de variáveis do solo, explorando uma ampla gama de abordagens, desde qualitativas – opiniões de especialistas e resultados de revisão bibliográfica – até metodologias semiquantitativas e quantitativas baseadas em estatísticas [30,31]. Há, no entanto, um consenso geral sobre a necessidade de se basear em abordagens rigorosas baseadas em métodos de seleção de características

Os métodos de seleção de características resultam em um subconjunto das características ou variáveis originais [32,33], selecionando variáveis “relevantes” que capturam a maior parte das informações sobre os processos e o funcionamento do sistema investigado, excluindo variáveis “irrelevantes” e/ou “redundantes”. Dentre as abordagens de seleção de características, a análise de componentes principais (ACP) tem sido utilizada principalmente [29]. Métodos supervisionados e técnicas regressivas, como análise discriminante e regressão por mínimos quadrados parciais, estão sendo progressivamente avaliados quanto à sua eficácia em diversos estudos [30,34,35].

A regressão por mínimos quadrados parciais (PLSR) é uma ferramenta poderosa usada em quimiometria [36] e provou ser muito versátil com aplicações em vários campos de pesquisa [32]. PLSR é um método supervisionado, relacionando as variações em uma (PLS1) ou múltiplas (PLS2) variáveis de resposta às variações de vários preditores; o método seleciona fatores ortogonais sucessivos, que maximizam a covariância entre preditores e variáveis de resposta. PLSR é particularmente útil quando o número de preditores é maior que o número de observações e/ou quando os preditores são altamente correlacionados, como frequentemente acontece com indicadores de solo. Embora o PLSR seja geralmente adotado para fins preditivos, um grande número de métodos para seleção de variáveis em PLSR foram propostos. Entre eles, métodos de filtro (vetores de peso de carga, coeficientes de regressão PLS e importância de variáveis em estatísticas de projeção - VIP) utilizam a saída do algoritmo PLS para identificar um subconjunto de variáveis importantes [32].

Em particular, as estatísticas VIP são capazes de resumir a contribuição dos preditores e das variáveis de resposta [37].

A análise de componentes principais (ACP) é um procedimento amplamente difundido, desenvolvido para resumir grandes conjuntos de dados de variáveis correlacionadas em um número reduzido de componentes que contêm a maior parte das informações originais [38]. Pesos ou cargas variáveis dos componentes retidos são úteis para identificar as variáveis que mais contribuem para cada CP selecionado e investigar suas relações. Correlações de variáveis e análises discriminantes passo a passo e canônicas (ADC e ACD) foram adotadas em estudos anteriores [15,33,39]. Em particular, a ADC permite a seleção de variáveis que melhor diferenciam tratamentos ou classes.

Em um estudo que teve como objetivo avaliar os efeitos de diferentes práticas de manejo na qualidade física do solo (QFS) em experimentos de campo de longo prazo, Castellini et al. [15] aplicaram análise de correlação, PCA e SDA para selecionar indicadores-chave de qualidade física do solo. Os resultados destacaram o papel complementar e suplementar dos três métodos aplicados e a importância de usar simultaneamente diferentes abordagens para obter uma compreensão completa dos processos investigados. De Paul Obade et al. [35], após uma comparação preliminar de diferentes abordagens multivariadas, usaram PLSR para definir um índice de qualidade do solo padronizado para diferentes condições de solo, vegetação natural ou florestas e diferentes manejos do solo, preparo convencional e plantio direto. Carbono orgânico do solo, densidade do solo, relação carbono-nitrogênio e condutividade elétrica foram identificados como as principais variáveis que influenciam o estado da qualidade do solo e o índice forneceu uma avaliação abrangente do manejo investigado.

Shukla et al. [24] utilizaram a análise fatorial para identificar indicadores apropriados de qualidade do solo sob diferentes práticas de uso e manejo da terra. O carbono orgânico do solo foi o atributo medido mais dominante; outros atributos-chave do solo foram principalmente indicadores físicos — ou seja, capacidade hídrica em campo, porosidade preenchida por ar, densidade do solo e pH. Pulido Moncada et al. [34] aplicaram árvores de decisão e avaliaram a qualidade do solo por meio da avaliação visual do solo em campo e de um número limitado de propriedades físicas e químicas do solo.

Independentemente da abordagem utilizada, pouca pesquisa está disponível sobre a comparação e o uso combinado de diferentes métodos estatísticos para seleção de indicadores de solo. Portanto, os principais objetivos deste estudo foram (i) identificar as variáveis mais adequadas para discriminar o estado do solo sob diferentes estratégias de manejo do solo e (ii) comparar o desempenho de três abordagens estatísticas multivariadas para selecionar indicadores de solo. Foram utilizados dados coletados em 2015 de um experimento de campo de longo prazo que investigou os efeitos de duas estratégias de manejo do solo em trigo duro (*Triticum durum* Desf.) em uma área mediterrânea. O desempenho do PCA foi comparado ao SDA e ao PLSR usando estatísticas VIP (PLS-VIP) para seleção de características.

2. Materiais e Métodos

2.1. Área de Estudo e Experimento de Campo de Longa Duração

O estudo foi realizado em um conjunto de dados coletados de um experimento de campo de longo prazo, realizado na fazenda experimental do Conselho de Pesquisa e Economia Agrícola (CREA-AA) no sul da Itália (Foggia, 41°27'03 N, 15°30'06 E).

O clima da área de estudo é classificado como “termomediterrâneo acentuado” [40], com temperaturas que podem cair abaixo de 0 °C no inverno e ultrapassar 40 °C no verão. A precipitação é distribuída de forma irregular ao longo do ano e concentra-se principalmente nos meses de inverno, com uma média anual de longo prazo de 550 mm [41]. O solo é argiloso de origem aluvial (42,7 e 29,6% de argila e areia, respectivamente), classificado como fino, mésico, cromoxereto típico [42].

O ensaio de monocultivo de trigo duro de sequeiro, iniciado em 2002 e ainda em andamento, compara os efeitos de duas estratégias de manejo do solo (ou seja, preparo mínimo do solo, MT, e plantio direto, NT) na resposta à produtividade do trigo duro e na fertilidade do solo. O preparo mínimo do solo consiste em um preparo do solo em duas camadas a 40 cm de profundidade (ou seja, uma combinação de escarificador e enxada rotativa) realizado no outono, antes da semeadura do trigo duro. O plantio direto consiste na semeadura direta do trigo duro após um tratamento químico de capina. Os tratamentos são organizados

em delineamento de blocos casualizados com três repetições e parcelas unitárias de 500 m². Para ambos os tratamentos, em setembro, a palha foi cortada em pedaços de 10 a 15 cm e espalhada de volta na parcela. A capina com glifosato foi realizada no início de novembro nas parcelas do NT. A semeadura foi realizada para ambos os tratamentos no final de novembro com uma semeadeira para semeadura direta, devidamente equipada com lâminas moldadas. Todas as outras técnicas agrônômicas (fertilização, controle de pragas e manejo de plantas daninhas durante o crescimento da cultura) foram realizadas uniformemente para os dois manejos de solo comparados. Na colheita, o rendimento foi medido em cada local do solo dentro de uma subárea de 1 × 1 m. Mais informações sobre o manejo da parcela podem ser encontradas em Castellini et al. [15].

2.2. Amostragem de solo e medições em laboratório

Amostras de solo não perturbadas foram coletadas na espiga de trigo (abril de 2015) dentro de cada unidade experimental do delineamento experimental RCBD em 4 locais sub-replicados, para um total de 24 observações. Amostras de solo foram coletadas a 0–0,20 m de profundidade. Os seguintes indicadores foram

quantificados nas amostras de solo: (i) indicadores químicos: carbono orgânico total e nitrogênio total (COT e N), carbono extraível por álcali (TEC), carbono de ácidos húmicos e fúlvicos (HA_FA), nitrogênio extraível por água e carbono orgânico (WEN e WEOC), fósforo disponível Olsen (P_Olsen), cátions trocáveis (Ca²⁺ e Mg²⁺) e condutividade elétrica (CE); (ii) indicadores físicos: textura, densidade aparente seca (DS), macroporosidade (PMAC), capacidade de ar (CA) e capacidade de campo relativa (CRC); (iii) indicador biológico: carbono da biomassa microbiana quantificado pelo método de fumigação-extração [43].

WEOC e WEN, que são indicadores de reservas orgânicas lábeis de C e N, foram extraídos de amostras de solo úmido de campo de acordo com o protocolo relatado em Armenise et al. [28]. TOC, TEC, HA_FA e N total foram quantificados em amostras secas e peneiradas de 2 mm, seguindo protocolos relatados em Ferrara et al. [17]. Em detalhes, para quantificação de TOC, amostras de solo foram moídas até um pó fino (0,5 mm) usando um moinho de bolas de ágata. TEC foi obtido por extração de 0,1 M NaOH + 0,1 M Na₄P₂O₇ a 65 °C por 48 h. Os ácidos húmico e fúlvico foram fracionados por acidificação a pH = 2,0 com H₂SO₄. A purificação de FA de substâncias não húmicas foi realizada por adsorção em colunas de polivinilpirrolidona. O carbono orgânico total em amostras de solo, bem como as frações de C em extratos alcalinos e as frações de C e N em extratos aquosos, foram quantificados por combustão seca usando um analisador TOC Vario Select (Elementar, Lomazzo, Alemanha), que realiza uma combustão catalítica da amostra em altas temperaturas em ambiente de ar. O N total foi quantificado de acordo com o procedimento de Kjeldahl.

Tanto o BD quanto alguns pontos da curva de retenção de água no solo, ou seja, a relação entre o conteúdo volumétrico de água no solo (θ) e a carga de pressão da água (h), foram determinados usando núcleos de solo não perturbados. Especificamente, anéis de aço inoxidável com bordas afiadas (diâmetro interno de 8 cm; altura de 5 cm) foram usados para determinar o BD do solo e a curva de retenção de água em altas cargas de pressão ($h \geq 120$ cm). Uma amostra de solo perturbado foi coletada perto dos pontos de coleta de amostra não perturbada para determinar a curva de retenção de água em baixas cargas de pressão ($h \leq 330$ cm). Os valores de θ foram determinados em cada núcleo de solo não perturbado por um aparelho de coluna de água suspensa [44] para valores de h variando de -5 a -120 cm, e em núcleos de solo reembalados pelo método de placa de pressão [45] para valores de h variando de -330 a -15.3 cm. A função de retenção de água no solo foi obtida ajustando os dados experimentais com o modelo de van Genuchten [47], e um conjunto de indicadores capacitivos que levam em conta a disponibilidade de água/ar foram estimados: macroporosidade (PMAC = $\theta_s - \theta_m$) (cm³ cm⁻³), capacidade de ar (AC = $\theta_s - \theta_{FC}$) (cm³ cm⁻³) e capacidade de campo relativa (RFC = θ_{FC}/θ_s) (adimensional), onde θ_s , θ_m , θ_{FC} são os conteúdos volumétricos de água correspondentes a uma carga de pressão de 0, θ_{10} e θ_{100} cm, respectivamente. A avaliação da qualidade física do solo (SPQ) foi realizada de acordo com classificações coletadas da literatura [14]. Em particular, o SPQ foi considerado ótimo quando $0,9 \leq BD \leq 1,2$ g cm⁻³, PMAC $\leq 0,07$ cm³ cm⁻³ AC $\leq 0,14$ cm³ cm⁻³.

e 0,6 \ddot{y} RFC \ddot{y} 0,7. Informações complementares sobre os indicadores descritos podem ser encontradas em Castellini et al. [15].

2.3. Análise de Dados

Estatísticas descritivas foram computadas para resumir as principais características da distribuição dos dados das variáveis de resposta a serem utilizadas na abordagem regressiva. Além disso, as variáveis foram testadas quanto à normalidade, utilizando os testes de Shapiro-Wilk e Kolmogorov-Smirnov, e quanto à heterocedasticidade pelo teste de Levene. A distribuição dos dados e a presença de heterocedasticidade também foram examinadas para as variáveis do solo.

O conjunto de vinte variáveis do solo foi inicialmente analisado por meio de uma análise de variância aninhada (ANOVA), considerando as réplicas dentro das parcelas como pseudo-réplicas. Em seguida, os dados foram analisados por meio da Análise de Componentes Principais (ACP), Análise Discriminante Stepwise (ADS) e Regressão por Mínimos Quadrados Parciais (RPM) com estatísticas de Importância da Variável para Projeção (VIP).

A ACP foi aplicada à matriz de correlação das variáveis do solo a fim de obter poucos componentes novos que explicassem a maior parte da variação das variáveis originais. Os componentes principais (CPs) que explicaram cumulativamente uma alta porcentagem da variância total e tiveram um autovalor maior que um (critério de Kaiser) foram retidos. Juntamente com o autovalor, a porcentagem de variação explicada pelo componente único foi levada em consideração, considerando o limite de 5% sugerido por Wander e Bollero [48]. Cargas variáveis foram examinadas. Dentro de cada CP, apenas cargas altamente ponderadas, definidas como tendo valores absolutos dentro de 10% da carga mais alta [49], foram consideradas e os sinais foram examinados para investigar as relações entre as variáveis selecionadas.

A ASD foi aplicada para identificar as variáveis que permitem a discriminação máxima entre as classes comparadas (tratamentos de manejo do solo). A estatística lambda de Wilks foi usada como medida multivariada de separabilidade [50]. O uso da ASD requer que um conjunto de suposições seja verificado, entre as quais a normalidade das distribuições de dados, a homocedasticidade e a redundância incompleta das variáveis consideradas. No entanto, um desvio moderado dessas suposições não afeta seriamente os resultados da análise, como demonstrado por uma ampla literatura sobre a aplicação da ASD [51–53]. A ASD foi realizada usando o algoritmo STEPWISE do procedimento STEPDISC do SAS/STAT [54]; o nível de significância para entrada e para permanência foi estabelecido em 0,05.

PCA e SDA foram primeiramente realizados separadamente no conjunto de variáveis químicas (mais carbono da biomassa microbiana) e variáveis físicas, e depois foram realizados em todo o conjunto de dados de indicadores.

O PLSR foi realizado em dados centrados na média e em escala de variância de preditores e variáveis de resposta. Neste estudo, os preditores foram representados por propriedades químicas, físicas e biológicas do solo; as variáveis de resposta foram o rendimento de grãos e o teor de proteína dos grãos, como indicadores integrados da resposta quantitativa e qualitativa da cultura a diferentes condições de manejo do solo. O número ótimo de fatores a serem retidos no modelo foi baseado na estatística da soma dos quadrados residual mínima prevista (PRESS) [54]. Neste estudo, ambas as abordagens foram usadas na execução do PLSR: PLS1 e PLS2 [55,56]. No PLS1, o PLSR foi realizado usando o rendimento de grãos ou o teor de proteína como variável de resposta única; no PLS2, o rendimento de grãos e o teor de proteína foram considerados simultaneamente como múltiplas variáveis de resposta.

Estatísticas VIP foram utilizadas para a seleção de variáveis [57]. Como a média dos quadrados das pontuações VIP é igual a 1 [58], apenas as variáveis com pontuação VIP maior que 1 são geralmente consideradas significativas. No entanto, limiares entre 0,83 e 1,21 também foram sugeridos [32]. Um valor de 0,8 também é considerado por Wold [59] para retenção ou exclusão de variáveis. As variáveis de solo com os maiores valores de VIP foram selecionadas. A regressão PLS foi realizada utilizando o procedimento PLS do SAS/STAT [54].

3. Resultados

3.1. Análise Estatística Preliminar

Análise estatística preliminar realizada sobre rendimento de grãos e teor de proteína, resposta variáveis no PLSR, apresentaram valores médios e medianos próximos e coeficiente de assimetria e curtose igual ou inferior a 0,5. Esses resultados foram confirmados por testes de normalidade que indicaram para ambas as variáveis um desvio não significativo da distribuição normal ($P = 0,5198$ e $P = 0,0970$ para o teste de Shapiro-Wilk, para rendimento de grãos e teor de proteína, respectivamente). As variâncias foram homogêneas ao longo dos tratamentos de manejo para ambas as variáveis de acordo com Teste de Levene ($P = 0,4128$ para rendimento de grãos e $P = 0,2432$ para teor de proteína). Valores médios de rendimento de grãos e teor de proteína foram respectivamente 5,247 Mg ha⁻¹ e 13,39 g 100 g⁻¹.

As distribuições das variáveis do solo não se desviaram significativamente (biomassa microbiana C, TOC, TEC, HA_FA, N, Olsen P, pH, K⁺, Mg²⁺, Na⁺, BD, AC, RFC, argila e areia) ou apenas ligeiramente desviado (EC, Ca²⁺, PMAC) da distribuição normal, exceto para WEOC e WEN. Homocedasticidade também foi observada na maior parte dos casos (exceto para WEN e PMAC). Por esse motivo, os dados foram analisados para todas as variáveis da escala original.

3.2. Análise de Variância

Diferentes estratégias de manejo do solo afetaram significativamente tanto a matéria orgânica total quanto a matéria orgânica total. concentrações de carbono extraível (COT e TEC) e nitrogênio extraível em água (WEN), com maior teor de TOC e TEC observado na camada superior do solo sob condições mais conservadoras manejo do solo (Tabela 1). Em particular, um valor médio de COT de 21,87 g kg⁻¹ foi registrado em plantio direto (PD) em comparação com 17,45 g kg⁻¹ em plantio mínimo (PM) gestão. Uma maior concentração significativa de potássio trocável (K⁺) também foi observadas em solos não cultivados (Tabela 1). Sob cultivo mínimo, maiores concentrações de WEN foram observados.

Tabela 1. Resultados da análise de variância aninhada realizada em parâmetros químicos e biológicos do solo.

Fonte de Variação	WEOC mg kg ⁻¹	WEN mg kg ⁻¹	C_biomassa mg kg ⁻¹	Índice g kg ⁻¹	TEC g kg ⁻¹	HA_FA g kg ⁻¹	N g kg ⁻¹	P_Olsen mg kg ⁻¹	pH	CE dS m ⁻¹	Ca ²⁺ mg kg ⁻¹	K ⁺ mg kg ⁻¹	Mg ²⁺ mg kg ⁻¹	Na ⁺ mg kg ⁻¹
Solo gerenciamento (significan) Pr(>F)	49.351 0,1732	25.186 0,0273 *	491,03 0,8461	19,66 0,0325 *	12.088 0,0415 *	6.782 0,3280	1.470 0,4441	54.045 0,4609	8.10 0,1674	0,139 0,6251	6880,9 0,9698	1043,89 0,0233 *	215,84 0,1663	29.122 0,1882
Plantio direto (NT)	70,44	10.11 b	509,8	21,87 a	13,63 a	6,23	1,52	60,72	8,07	0,14	6880,3	1120,18 a	211,65	27,87
Mínimo lavoura (MT)	28,26	40,26 a	472,3	17h45	10,54 b	7,34	1,42	47,37	8,14	0,13	6881,6	967,60 b	220,03	30,37

* e ** indicam, respectivamente, diferenças em $P \leq 0,05$ e $P \leq 0,01$. Médias seguidas por letras diferentes são significativamente diferentes de acordo com o teste SNK ($P = 0,05$). Pr(> F) indica o valor de probabilidade (valor-p) para determinar se a hipótese nula deve ser rejeitada. WEOC = carbono orgânico extraível da água; WEN = nitrogênio extraível da água; C_biomass = carbono da biomassa microbiana; TOC = carbono orgânico total; TEC = carbono extraível por álcali; HA_FA = carbono de ácido húmico e fúlvico; N = nitrogênio total; P_Olsen = fósforo disponível de Olsen; CE = condutividade elétrica; Ca²⁺, Mg²⁺, Na⁺, K⁺ = cátions trocáveis. (n = 24).

A gestão do solo também afetou as propriedades físicas com um volume significativamente maior densidade (DB) em solos não cultivados e, como consequência, menor capacidade de ar (CA), indicando tendência à compactação do solo (Tabela 2). A macroporosidade (PMAC) confirmou essa tendência, embora não tenham sido registradas diferenças significativas. Capacidade de campo relativa (RFC), que fornece um relato do equilíbrio entre a capacidade de água e a capacidade de ar do solo (em outras palavras, palavras, é um índice da importância relativa dos meso-microporos para a porosidade total), foi maior em NT do que em MT, sugerindo grandes riscos potenciais de condições anaeróbicas devido a redução da presença de ar na porosidade do solo.

Tabela 2. Resultados da análise de variância aninhada realizada em parâmetros físicos e hidrológicos do solo.

Fonte de variação	BD g cm ³	PMAC cm ³ cm ³	CA cm ³ cm ³	RFC -	Argila g 100 g ⁻¹	Areia g 100 g ⁻¹
Gestão do solo (média)	0,96729	0,03329	0,08473	0,8151	48,15	11,03
Pr(>F)	0,0316 *	0,0992	0,0370 *	0,0302 *	0,1854	0,5042
Plantio direto (PD)	1,04516 a	0,00890	0,04112 b	0,90814 a	45,44	11,41
Cultivo mínimo (MT)	0,90240 b	0,05362	0,12107 a	0,73764 b	50,85	10,65

* e ** indicam, respectivamente, diferenças em P \leq 0,05 e P \leq 0,01. Médias seguidas por letras diferentes são significativamente diferentes de acordo ao teste SNK (P = 0,05). Pr(> F) indica o valor de probabilidade (valor-p) para determinar se a hipótese nula deve ser rejeitada. BD = granel seco densidade; PMAC = macroporosidade; AC = capacidade de ar; RFC = capacidade de campo relativa. (n = 22 para BD, PMAC, AC, RFC. n = 24 para argila e areia).

Segundo Reynolds et al. [14], foram observados valores ótimos e intermediários sob MT para densidade aparente (0,9–1,2 g cm³, alcance ideal), capacidade de ar (0,10–0,14 cm³ cm³, faixa intermediária) e macroporosidade (0,04–0,07 cm³ cm³, faixa intermediária) e valores ligeiramente acima do limite ideal para RFC (0,6–0,7, intervalo ideal). Exceto para BD, os valores registrados em NT foram todos indicativos de condições de aeração limitadas.

A iteração a longo prazo das diferentes estratégias de gestão afetou ligeiramente os grãos resposta de rendimento (P = 0,0603), com valores médios de 5,45 e 5,04 Mg ha⁻¹ registrados sob plantio direto e preparo mínimo do solo, respectivamente. Nenhum efeito significativo do manejo comparado foi observado em vez disso para o conteúdo de proteína dos grãos.

3.3. Análise de Componentes Principais

A PCA foi realizada inicialmente separadamente no conjunto de variáveis químicas e físicas, e então realizado em todo o conjunto de dados.

Na análise dos indicadores químicos, os três primeiros componentes (CPs) explicaram cerca de 65,16% da variância total, enquanto na análise dos indicadores físicos, o primeiro dois PCs foram capazes de explicar 80,12% da variância total (Tabela 3).

Tabela 3. Autovalores e variância explicados pelos cinco primeiros componentes principais (CPs) da análise realizada no conjunto de (a) indicadores químicos e biológicos (14 variáveis) e (b) indicadores físicos (6 variáveis).

(a) Autovalores da Matriz de Correlação: Total =14 Média = 1					(b) Autovalores da Matriz de Correlação: Total =6 Média = 1				
Proporção de diferença de autovalor cumulativa					Proporção de diferença de autovalor cumulativa				
1	5,1000	2,6932	0,3643	0,3643	1	3,7885	2,7698	0,6314	0,6314
2	2,4068	0,7911	0,1719	0,5362	2	1,0187	0,1054	0,1698	0,8012
3	1,6158	0,4237	0,1154	0,6516	3	0,9133	0,6757	0,1522	0,9534
4	1,1921	0,2644	0,0851	0,7368	4	0,2376	0,1971	0,0396	0,993
5	0,9278	0,2123	0,0663	0,803	5	0,0404	0,0389	0,0067	0,9998

Os gráficos de pontuação dos dois primeiros componentes mostraram que tanto os componentes químicos como os físicos variáveis foram capazes de discriminar os diferentes manejos do solo comparados (Figura 1). A inspeção das cargas dos primeiros PCs destacou que os solos sem cultivo (SD) eram caracterizado por um maior TOC e TEC, juntamente com conteúdo de K+ trocável, e por um maior densidade aparente e RFC, enquanto menor PMAC e AC foram detectados.

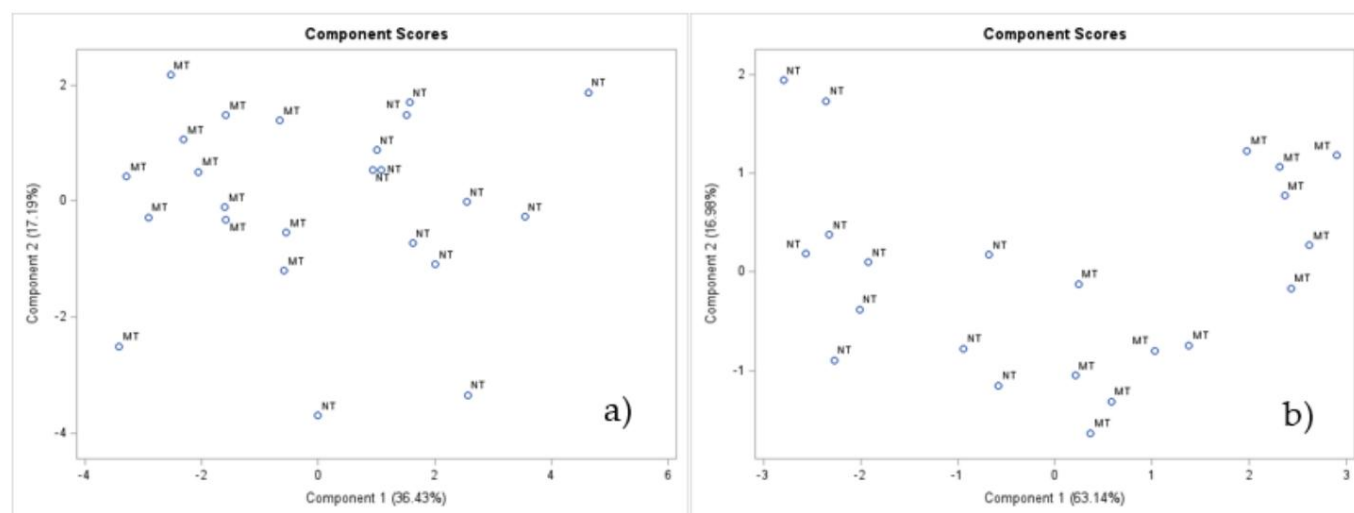


Figura 1: Gráficos de pontuação dos dois primeiros componentes principais extraídos na análise realizada no conjunto de substâncias químicas (a) e indicadores físicos (b) do solo. MT = cultivo mínimo; NT = plântio direto.

A análise de todo o conjunto de indicadores físicos e químicos confirmaram esses resultados com uma discriminação de tratamento ainda mais clara e o mesmo acontece com as variáveis físicas e químicas. A análise de todo o conjunto de indicadores físicos e químicos confirmaram esses resultados com uma discriminação de tratamento ainda mais clara e o mesmo acontece com as variáveis físicas e químicas. A análise de todo o conjunto de indicadores físicos e químicos confirmaram esses resultados com uma discriminação de tratamento ainda mais clara e o mesmo acontece com as variáveis físicas e químicas.

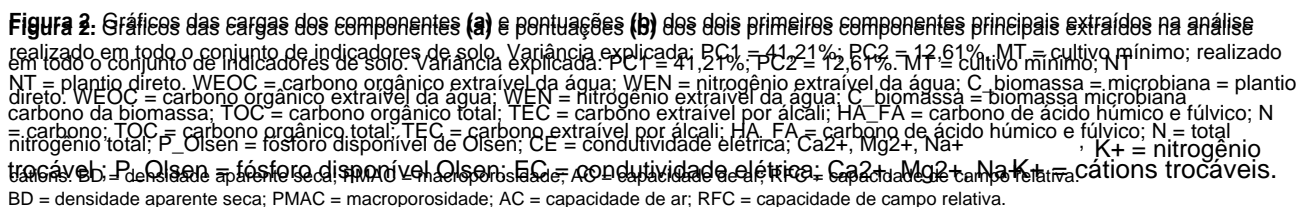
Tabela 4. Autovalores e variância explicados pelos cinco primeiros componentes principais (CPs) da análise realizada em todo o conjunto de dados de indicadores de solo (20 variáveis).

Autovalores da Matriz de Correlação: Total						
Autovalores da Matriz de Correlação: Total						
= 20 Média = 1						
		Autovalor	Autovalor	Proporção de	Proporção de	Cumulativo
				Variação	Variação	
1	1	8,2424	8,2424	5,7211	5,7211	0,4121
2	2	2,5213	2,5213	0,3911	0,3911	0,5332
3	3	2,1302	2,1302	0,5153	0,5153	0,6447
4	4	1,6149	1,6149	0,4059	0,4059	0,7254
5	5	1,2089	1,2089	0,2821	0,2821	0,7859

Tabela 5. Cargas variáveis dos quatro primeiros componentes da análise realizada em todo o conjunto de dados. Os valores são multiplicados por 100 e arredondado para o número inteiro mais próximo. Variância explicada: PC1 = 41,21%; PC2 = 12,61%; PC3 = 10,65%; PC4 = 8,07%.

Tabela 5. Cargas variáveis dos quatro primeiros componentes da análise realizada em todo o conjunto de dados.									
Os valores são multiplicados por 100 e arredondados para o número inteiro mais próximo. Variância explicada: PC1 = 41,21%; PC2 = 12,61%; PC3 = 10,65%; PC4 = 8,07%.									
Fator 1	Fator 2	Fator 3	Fator 4	Fator 1	Fator 2	Fator 3	Fator 4	Fator 1	Fator 2
WEOC	54								
WEN	77								
C_biomassa	32								
Índice	91								
TEC	93								
HA_FA	36								
N	36								
P_Olsen	33								
pH	70								
CE	50								
Ca2+									
K+	2								
Mg2+	82								
Na+	46								
BD	52								
PMAC	84								
CA	91								
Areia	92								
argilosa RFC	23								

* indica a significância das cargas variáveis. O sinal das cargas variáveis indica a correlação positiva ou negativa entre as variáveis e o componente principal. WEOC = carbono orgânico extraível com água; WEN = nitrogênio extraível com água. C_biomassa = carbono da biomassa microbiana; TOC = carbono orgânico total; TEC = carbono extraível por álcool; HA_FA = ácido húmico e fúlvico; N = nitrogênio total; P_Olsen = fósforo disponível de Olsen; CE = condutividade elétrica; Ca2+, Mg2+, Na+, K+ = trocáveis cations. BD = densidade aparente seca; PMAC = macroporosidade; AC = capacidade de ar; RFC = capacidade de campo relativa.



Em detalhe, os quatro primeiros PCs explicaram cumulativamente cerca de 72,54% da variância total (Tabela 4). No primeiro PC (41,21% da variância total), as variáveis altamente ponderadas foram TOC (Tabela 4). No primeiro PC (41,21% da variância total), as variáveis altamente ponderadas foram e TEC entre as variáveis químicas, enquanto RFC e AC, seguidos por PMAC, para o hy- TOC e TEC entre as variáveis químicas, enquanto RFC e AC, seguidos por PMAC, para o hy- Parametros hidrologicos do solo (Tabela 5). Um pouco abaixo do limite de 10% dos parametros hidrologicos mais altos fator de carga [50], havia K+ trocável e densidade em massa e, em menor extensão, fator de carga [50], havia K+ trocável e densidade em massa e, em menor extensão, WEN. No segundo PC (12,61%), o P disponível e os ácidos húmicos e fúlvicos apresentaram o alto WEN. Nas cargas mais elevadas, com carbono da biomassa microbiana e pH ligeiramente abaixo do limite. Nas cargas mais elevadas, com carbono da biomassa microbiana e pH ligeiramente abaixo do limite. terceiro (10,65%) e quarto (8,07%) componentes. Ca²⁺ e N.trocáveis foram selecionados, o terceiro (10,65%) e quarto (8,07%) componentes, Ca²⁺ e N.trocáveis foram selecionados, respectivamente (Tabela 5). A inspeção desses resultados mostrou que o primeiro PC resumiu principais resultados da análise de variância. O segundo componente destacou, em vez disso, os resultados da análise de variância. O segundo componente destacou, em vez disso, os principais resultados da análise de variância. O papel dos nutrientes disponíveis (P) e de algumas frações de carbono (ácidos húmicos e fúlvicos C e papel dos nutrientes disponíveis (P) e de algumas frações de carbono (ácidos húmicos e fúlvicos C e papel da biomassa microbiana C), adicionando, desta forma, mais elementos para explicar as diferenças observadas nas condições experimentais.

A SDA foi realizada primeiramente separadamente no conjunto de variáveis químicas e físicas, e então realizado em todo o conjunto de dados.

Variáveis que permitem a discriminação máxima entre os tratamentos para a análise realizada no conjunto de dados de indicadores químicos e biológicos foram TOC ($P < 0,0001$) e HA_FA ($P = 0,0004$), seguido em menor extensão por Olsen P ($P = 0,0111$), EC ($P = 0,0258$), WEOC ($P = 0,0397$) e TEC ($P = 0,05$). Na análise realizada sobre o conjunto de indicadores físicos, as variáveis selecionadas foram RFC ($P < 0,0001$) e argila ($P = 0,037$).

Por fim, a análise realizada em todo o conjunto de dados sintetizou os resultados obtidos, selecionando TOC ($P < 0,0001$), RFC ($P < 0,0001$) e WEOC ($P = 0,003$) como variáveis que possibilitaram discriminação máxima entre tratamentos (Tabela 6).

Tabela 6. Seleção resumida do procedimento STEPDISC realizado em todo o conjunto de dados de análises químicas, físicas e biológicas. indicadores.

Etapa	Número inserido Removido Parcialmente			Valor F	P > F	Wilks'	Pr <	Média	Pr >
	Em		R-Quadrado			Lambda	Lambda ao quadrado	Pode Corr	ASCC
1	1	Índice	0,8297	97,46	<0,0001	0,17026571	<0,0001	0,82973429	<0,0001
2	2	RFC	0,5631	24,49	<0,0001	0,07439379	<0,0001	0,92560621	<0,0001
3	3	WEOC	0,3941	11,71	0,003	0,0450772	<0,0001	0,9549228	<0,0001

Tanto a PCA quanto a SDA, bem como a análise de variância univariada, retornaram TOC e RFC entre as variáveis mais influentes tanto no conjunto de dados químicos quanto físicos analisados separadamente, bem como em todo o conjunto de dados. Além disso, o WEOC foi selecionado entre as variáveis mais discriminantes no SDA realizado em todo o conjunto de dados de indicadores de solo. Esses achados estão de acordo com os resultados relatados por estudos anteriores na seleção dos indicadores mais relevantes para avaliar o estado da qualidade do solo.

3.5. Regressão de Mínimos Quadrados Parciais e Estatísticas VIP

O PLSR foi aplicado para obter mais informações na seleção das variáveis mais informativas para descrever os efeitos das duas práticas de manejo do solo comparadas.

O método foi aplicado a todo o conjunto de indicadores de solo e considerando a produtividade de grãos e conteúdo de proteína de grãos como variáveis de resposta única (PLS1) e como variáveis de resposta múltipla variáveis (PLS2).

Quando o rendimento de grãos de trigo foi considerado como variável de resposta única, as duas primeiras fatores foram responsáveis cumulativamente por 47,20% da variância total nos preditores e 48,17% no variável de resposta. Os perfis de coeficiente VIP (Figura 3a) mostraram que os valores mais altos foram registrados para Olsen P e PMAC, seguidos de TOC, TEC, pH e Mg²⁺ (para o produto químico variáveis) e por argila, RFC e AC (para as variáveis físicas).

A utilização do teor de proteína dos grãos como variável de resposta modificou ainda mais a classificação dos indicadores selecionados. Os dois primeiros fatores representaram cumulativamente 49,54% do total variância nos preditores e 23,59% na resposta. Os maiores valores da estatística PLS-VIP foram registrados para N e Olsen P, seguidos por WEN e areia, indicando assim a contribuição de macronutrientes, com particular atenção ao N (forma total e lábil), que afecta os grãos concentração de proteína (Figura 3b). Além disso, TOC, TEC, Ca²⁺ e pH apresentaram valores importantes contribuição.

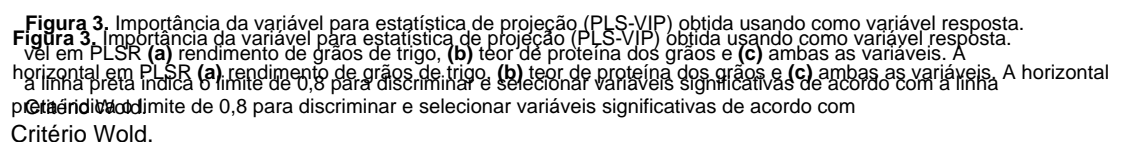
A inspeção de ambos os perfis VIP (Figura 3a,b) sublinhou o papel do físico indicadores de solo (PMAC, AC e RFC), que foram selecionados como variáveis importantes em ambos análises.

Por fim, o PLSR foi realizado considerando simultaneamente, como variáveis de resposta, o rendimento de grãos e o teor de proteína (PLS2). Os dois primeiros fatores extraídos foram responsáveis cumulativamente para 48,56% da variação total nos preditores e 27,16% na variável de resposta. A partir da inspeção do perfil VIP (Figura 3c), o papel dos nutrientes disponíveis (Olsen P e, secundário, WEN) e variáveis hidrológicas (PMAC, AC e RFC) na explicação da planta a resposta emergiu novamente. Essas variáveis foram seguidas pelos teores de carbono do solo (COT e TEC), pH e cátions trocáveis.

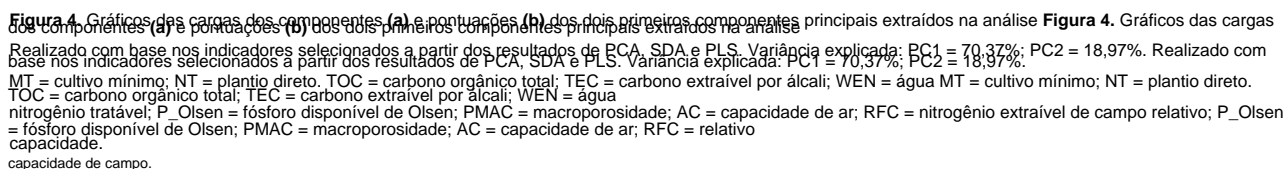
A inclusão de uma variável de resposta no PLSR — rendimento, conteúdo proteico ou ambos — modificou a classificação dos indicadores de solo selecionados, dando maior ênfase aos macronutrientes das plantas (P e N disponíveis), principalmente quando o teor de proteína dos grãos foi utilizado como variável dependente. Esses resultados ressaltam a importância do uso de abordagens combinadas explorar os dados e interpretar o comportamento do sistema investigado.

Finalmente, para um propósito exploratório adicional, a ACP foi realizada nas variáveis mais frequentemente selecionados com o uso combinado de SDA, PCA e PLSR, nomeadamente TOC, TEC, Olsen P, WEN, RFC, PMAC e AC. A Figura 4 mostra que as sete variáveis foram capazes de discriminar os tratamentos em investigação. Além disso, observou-se uma clara melhoria na

Appl. Sci. 2021, 11, x PARA REVISÃO POR PARES



12 de 17



O manejo do solo sem plantio direto a longo prazo também melhorou a matéria orgânica total e os extratos totais de carbono (COT e TEC). O comportamento observado está de acordo com vários estudos [19,63,64], uma vez que a redução da perturbação do solo reduz a rotatividade dos agregados do solo favorecendo o acúmulo e a estabilização da matéria orgânica nos micro e macroagregados, levando a um ganho líquido de carbono no solo. Valores significativamente diferentes de C e N da biomassa microbiana, densidade aparente, condutividade hidráulica e peso médio de Agregados de solo também foram observados por Sharma et al. [49] em plantio direto e convencional em um experimento de campo de longo prazo. Laudicina et al. [65], comparando o efeito de diferentes sistemas de cultivo (trigo/trigo e trigo/feijão) e manejos de preparo do solo mais utilizados (convencional, dupla camada e plantio direto), em um experimento de campo de longo prazo sobre C orgânico do solo (C orgânico total e extraível, C da biomassa microbiana, respiração basal), observados que o manejo da lavoura afetou o C orgânico do solo armazenado nos primeiros 15 cm de solo mais do que

sistema de cultivo. O plantio direto causou um aumento de 3,6 Mg ha⁻¹ no conteúdo de C em uma rotação trigo/faba e um aumento de 5,6 Mg ha⁻¹ na monocultura de trigo após 19 anos [65]. Um maior conteúdo de K⁺ trocável também foi observado no plantio direto. A ACP e a ASD confirmaram e resumiram os resultados da análise de variância, mas também destacaram o papel das frações de carbono orgânico - ácidos húmicos e fúlvicos e WEOC - e do P disponível como principais fontes de variabilidade na descrição dos dados (ACP) e como as variáveis que mais discriminaram os tratamentos comparados (SDA). Em um estudo que avaliou a adequação de diferentes frações de C lábeis como indicadores de qualidade do solo, Bongiorno et al. [66] descobriram que o WEOC (C orgânico dissolvido) não era sensível ao preparo do solo, ao contrário do C orgânico particulado; em qualquer caso, o conteúdo WEOC depende muito das condições ambientais e do tempo de amostragem do solo [66]. De acordo com o presente estudo, López-Fando e Pardo [67] mediram concentrações mais altas de K⁺ trocável e P disponível sob plantio direto em comparação ao preparo mínimo, a uma profundidade do solo de 0–20 cm. Da mesma forma, Martín-Rueda et al. [68] encontraram maiores concentrações de K e P disponíveis para plantas na superfície do solo (0–20 cm) sob plantio direto em comparação ao sistema de preparo mínimo, mas nenhuma diferença foi observada entre os dois sistemas de manejo do solo em maiores profundidades do solo. O acúmulo de P, K e outros nutrientes disponíveis nas camadas superficiais do solo sob plantio direto é geralmente atribuído à decomposição de matéria orgânica (que é mais abundante em solos sem plantio direto) e ao acúmulo de fertilizantes minerais na camada superficial do solo [69]. O sistema de preparo também influencia as relações que ocorrem entre raízes de plantas, solo e microrganismos no nível da rizosfera. Após plantio direto de longo prazo, uma maior atividade de fosfatase alcalina e fosfatase ácida foi medida por Balota et al. [70]. Essas duas enzimas, que podem ser liberadas tanto pelas raízes das plantas quanto pelos microrganismos do solo, estão envolvidas na liberação de P lábil dos reservatórios orgânicos. Além disso, os ácidos orgânicos exsudados pelas raízes das plantas e/ou liberados pela decomposição da matéria orgânica podem competir com o P pelos sítios de ligação nas partículas do solo, aumentando assim a disponibilidade de P [69].

Finalmente, o PLSR, ao considerar simultaneamente indicadores do solo e a resposta da planta (rendimento de grãos, teor de proteína ou ambos), selecionou como variáveis importantes os nutrientes minerais (P disponível e N total e extraível de água), particularmente quando o teor de proteína do grão foi usado como variável dependente, juntamente com indicadores de qualidade física do solo (PMAC, AC, RFC), pH e cátions trocáveis. Nesse sentido, as estatísticas VIP, sendo uma soma ponderada dos quadrados dos coeficientes X-score do PLS para os componentes retidos, com os pesos calculados a partir da quantidade de variância da variável dependente (Y) explicada por cada componente retido [33,37], foram capazes de resumir a contribuição dos preditores e das variáveis de resposta. Assim, a contribuição dos nutrientes minerais na determinação da resposta da planta também foi destacada. Os resultados tanto da PCA quanto da SDA, bem como da análise de variância univariada, retornaram o TOC e o RFC entre as variáveis mais influentes tanto no conjunto de indicadores químicos e físicos analisados separadamente quanto em todo o conjunto de dados. Estudos anteriores destacaram o papel do RFC entre os indicadores do solo na avaliação da qualidade física do solo. Esta variável foi capaz de resumir parte das informações fornecidas por AC e PMAC e, apoiada também por BD e conteúdo de água disponível para as plantas, mostrou a maior capacidade de discriminação das estratégias de manejo do solo e resíduos de culturas comparadas [15]. A seleção do COT é importante porque é uma entrada necessária para indicadores de qualidade estrutural do solo [71]. Esta variável também é indicativa da qualidade química do solo, sendo positivamente relacionada com a CTC do solo e a capacidade de retenção de nutrientes. Além disso, o COT também pode ser indicativo da qualidade biológica do solo, devido ao seu papel fundamental na manutenção das relações tróficas do solo e no estímulo à atividade microbiana das plantas e do solo. Shukla et al. [24], selecionando indicadores-chave do solo por meio de análise fatorial, concluíram que o COT foi o atributo do solo medido mais dominante como indicador de qualidade do solo para as duas profundidades do solo investigadas e sugeriram seu uso para monitorar mudanças na qualidade do solo [24,34].

De modo geral, a inspeção dos perfis VIP, juntamente com os resultados de todos os métodos comparados, destacou o papel dos indicadores físicos do solo (PMAC, AC e RFC), que foram selecionados como variáveis importantes em todas as análises realizadas. Isso reforça a ideia de que tais indicadores capacitivos do solo (por exemplo, RFC) podem ser sugeridos para diversas aplicações práticas

incluindo, por exemplo, a estimativa da taxa ótima de emendas em laboratório, antes do uso em campo.

Na avaliação de modificações na qualidade do solo em decorrência do manejo agrônomo, o número e o tipo de indicadores considerados estão estritamente relacionados ao objetivo do estudo e à escala espacial investigada [72]. Nossa pesquisa foi realizada em escala de campo e teve um objetivo metodológico e exploratório, focado na comparação da força e da capacidade de diferentes abordagens estatísticas para extrair informações cruciais de dados de solo e obter uma melhor compreensão do sistema investigado. De qualquer forma, a abordagem combinada descrita neste estudo pode ser aplicada à análise de diferentes conjuntos de dados e condições, incluindo também diferentes escalas espaciais.

5. Conclusões

Este estudo metodológico demonstra a eficácia do uso de métodos de seleção de variáveis para resumir as informações derivadas de conjuntos de dados multivariados e aprimorar a compreensão e a interpretabilidade do sistema investigado. Os resultados também destacam a importância do uso simultâneo de diferentes abordagens, pois elas podem fornecer informações distintas e complementares. As abordagens estatísticas comparadas forneceram resultados distintos em termos de seleção e classificação das variáveis selecionadas. A presença de uma variável resposta, na técnica regressiva, impulsionou significativamente o processo de seleção de características. Em particular, a inclusão do rendimento ou do teor de proteína como variáveis de resposta no PLSR modificou a classificação dos indicadores de solo selecionados, dando maior ênfase aos nutrientes das plantas, principalmente quando o teor de proteína dos grãos foi considerado. As variáveis selecionadas com maior frequência com o uso combinado dos três métodos (COT, TEC, Olsen P, WEN, RFC, PMAC, AC) foram capazes de fornecer uma discriminação clara entre os tratamentos comparados, conforme demonstrado pela ACP realizada no conjunto de dados reduzido. Os resultados, por fim, enfatizam o papel dos conjuntos de dados plurianuais, que são ferramentas inestimáveis para explorar os benefícios e os limites de diferentes metodologias.

Contribuições dos autores: Conceitualização, AMS e MC; metodologia, AMS; análise formal, AMS; investigação, MC, RR e AMS; curadoria de dados, AMS, MC, MD, RR e CEG; redação — preparação do rascunho original, AMS e MC; redação — revisão e edição, AMS, MC, MD, RR e CEG. Todos os autores leram e concordaram com a versão publicada do manuscrito.

Financiamento: Os autores gostariam de agradecer à UE e ao MIUR pelo financiamento da presente contribuição metodológica, no âmbito do consórcio internacional colaborativo DESERT, financiado pela Chamada ERA-NET Cofund WaterWorks2014; esta ERA-NET é parte integrante das Atividades Conjuntas de 2015 desenvolvidas pela Iniciativa do Programa Conjunto Desafios da Água para um Mundo em Mudança (Water JPI).

O trabalho também foi apoiado pelos projetos “BIOTILLAGE, Abordagens inovadoras para melhorar o desempenho ambiental e produtivo dos sistemas de plantio direto de cereais”, financiados pelo PSR-Basilicata 2007–2013 e Fundação PRIMA, convocatória 2019-Seção 1—GA n.1912 “Abordagens participativas baseadas em pesquisa para a adoção da Agricultura de Conservação na Área Mediterrânea — projeto CAMA”.

Declaração de Consentimento Informado: Não aplicável.

Declaração de disponibilidade de dados: Os dados apresentados neste estudo estão disponíveis mediante solicitação ao autor correspondente.

Agradecimentos: Os autores agradecem a Carolina Vitti, Luisa Giglio e Marcello Mastrangelo pelo trabalho especializado em análises laboratoriais. Além disso, agradecem a Sabrina Moscelli, Vittorio Vonella e Franco Fornaro pela coleta de dados e condução dos ensaios agrônômicos.

Conflitos de interesse: Os autores declaram não haver conflito de interesse.

Abreviações

CA	capacidade de ar
BD	densidade aparente seca
C_biomassa	carbono da biomassa microbiana
HA_FA	carbono de ácido húmico e fúlvico
MT	cultivo mínimo
NT	plantio direto
N	nitrogênio total
PCA	análise de componentes principais
Regressão de mínimos quadrados parciais PLSR	
Macroporosidade PMAC	
P_Olsen	Olsen fósforo disponível
Capacidade de campo relativa RFC	
Análise discriminante passo a passo SDA	
Qualidade física do solo SPQ	
IQS	índices de qualidade do solo
Carbono extraível alcalino TEC	
TOC	carbono orgânico total
Importância das variáveis VIP e PLS-VIP para projeção	
	nitrogênio extraível com água
WEOC	carbono orgânico extraível em água WEN

Referências

- Palm, C.; Sanchez, P.; Ahamed, S.; Awiti, A. Solos: Uma perspectiva contemporânea. *Annu. Rev. Environ. Resour.* **2007**, *32*, 99–129. [\[Referência Cruzada\]](#)
- Lozano-Baez, SE; Domínguez-Haydar, Y.; Meli, P.; van Meervel, I.; Vásquez, KV; Castellini, M. Principais lacunas no monitoramento do solo durante a Restauração Florestal na Colômbia. *Restor. Ecol.* **2021**. Manuscrito de Autor Aceito. [\[CrossRef\]](#)
- Lal, R. Impactos do sequestro de carbono no solo nas mudanças climáticas globais e na segurança alimentar. *Science* **2004**, *304*, 1623–1627. [\[CrossRef\]](#)
- Lal, R. Restaurando a qualidade do solo para mitigar sua degradação. *Sustentabilidade* **2015**, *7*, 5875–5895. [\[CrossRef\]](#)
- Zornoza, R.; Acosta, JA; Bastida, F.; Domínguez, SG; Toledo, DM; Faz, A. Identificação de indicadores sensíveis para avaliar o inter-relação entre qualidade do solo, práticas de manejo e saúde humana. *Soil* **2015**, *1*, 173–185. [\[CrossRef\]](#)
- Buehmann, EK; Bom dia, G.; Baic, Z.; Creamerb, RE; De Deyn, G.; de Goede, R.; Fleskens, L.; Geissend, V.; Kuypers, TW; Mäder, P.; et al. Qualidade do solo — Uma revisão crítica. *Soil Biol. Biochem.* **2018**, *120*, 105–125. [\[CrossRef\]](#)
- Svoray, T.; Hassid, I.; Atkinson, PM; Moebius-Clune, BN; van Es, HM Mapeamento da saúde do solo em grandes áreas de importância agrícola. *Soil Sci. Soc. Am. J.* **2015**, *79*, 1420–1434. [\[CrossRef\]](#)
- Castellini, M.; Stellacci, AM; Di Prima, S.; Iovino, M.; Bagarello, V. Metodologia aprimorada de execução de beerkan para avaliar o impacto da água Efeitos na infiltração e nas propriedades hidráulicas de um solo franco sob plantio direto e convencional. *Soil Sci. Soc. Am. J.* **2021**. Manuscrito de autor aceito. [\[CrossRef\]](#)
- Castellini, M.; Stellacci, AM; Sisto, D.; Iovino, M. O impacto mecânico da água afetou a qualidade física do solo de um solo franco sob cultivo mínimo e plantio direto: uma avaliação usando execuções multi-altura de Beerkan e o procedimento BEST. *Land* **2021**, *10*, 195. [\[Referência Cruzada\]](#)
- Nouwakpo, SK; Song, J.; Gonzalez, JM Avaliação da estabilidade estrutural do solo com leito fluidizado, estabilidade de agregados e Simulação de precipitação em sistemas de preparo do solo e rotação de culturas de longo prazo. *Soil Tillage Res.* **2018**, *178*, 65–71. [\[CrossRef\]](#)
- Castellini, M.; Di Prima, S.; Moret-Fernández, D.; Lassabatere, L. Métodos de medição rápidos e precisos para determinar o solo Propriedades hidráulicas: Uma revisão. *J. Hydrol. Hydromech.* **2021**, *69*, 1–19. [\[CrossRef\]](#)
- Pranagal, J.; Wozniak, A. 30 anos de monocultura de trigo e cultivo reduzido e condição física do Phaeozem Rendzic. *Agric. Gestão de Águas.* **2021**, *243*, 106408. [\[CrossRef\]](#)
- Castellini, M.; Vonella, AV; Ventrella, D.; Rinaldi, M.; Baiamonte, G. Determinação das propriedades hidráulicas do solo usando infiltrômetro técnicas: Uma avaliação da variabilidade temporal em um experimento de longo prazo sob manejo de solo com plantio direto e mínimo. *Sustentabilidade* **2020**, *12*, 5019. [\[CrossRef\]](#)
- Reynolds, WD; Drury, CF; Tan, CS; Fox, CA; Yang, XM Uso de indicadores e características de função de volume de poros para quantificar a qualidade física do solo. *Geoderma* **2009**, *152*, 252–263. [\[CrossRef\]](#)
- Castellini, M.; Stellacci, AM; Barca, E.; Iovino, M. Aplicação de técnicas de análise multivariada para seleção de parâmetros físicos do solo Indicadores de qualidade: Um estudo de caso em experimentos de campo de longo prazo na Apúlia (sul da Itália). *Soil Sci. Soc. Am. J.* **2019**, *83*, 707–720. [\[Referência Cruzada\]](#)
- Peigné, J.; Ball, BC; Roger-Estrade, J.; David, C. A lavoura de conservação é adequada para a agricultura orgânica? Uma revisão. *Gerenciamento do Uso do Solo.* **2007**, *23*, 129–144. [\[Referência Cruzada\]](#)

17. Ferrara, RM; Mazza, G.; Muschitiello, C.; Castellini, M.; Stellacci, AM; Navarro, A.; Lagomarsino, A.; Vitti, C.; Rossi, R.; Rana, G. Efeitos de curto prazo da conversão para plantio direto na respiração e nas propriedades físico-químicas do solo: Um estudo de caso em um sistema de cultivo de trigo em ambiente semiseco. **2017**, *1*, 47–58.
18. Devine, S.; Markewitz, D.; Hendrix, P.; Coleman, D. Agregados do solo e matéria orgânica associada sob cultivo convencional, plantio direto e sucessão florestal após três décadas. *PLoS ONE* **2014**, *9*, e84988. [\[CrossRef\]](#)
19. Six, J.; Paustian, K. Dinâmica de agregados e MOS em sistemas de plantio direto e convencional. *Soil Sci. Soc. Am. J.* **1999**, *63*, 1350–1358. [\[Referência Cruzada\]](#)
20. Six, J.; Paustian, K. Rotatividade de macroagregados do solo e formação de microagregados: um mecanismo para sequestro de C. sob agricultura de plantio direto. *Soil Biol. Biochem.* **2000**, *32*, 2099–2103. [\[CrossRef\]](#)
21. Reynolds, WD; Schoenau, JJ; Drury, CF Benefícios científicos e práticos de estudos de campo de longo prazo. *Can. J. Soil Sci.* **2014**, *94*, 253–254. [\[Referência Cruzada\]](#)
22. Diácono, M.; Fiore, A.; Farina, R.; Canali, S.; Di Bene, C.; Testani, E.; Montemurro, F. Estratégias agroecológicas combinadas para adaptação de sistemas hortícolas orgânicos às mudanças climáticas no ambiente mediterrâneo. *Italiano. J. Agron.* **2016**, *11*, 85–91. [\[Referência Cruzada\]](#)
23. Diácono, M.; Persiani, A.; Canali, S.; Montemurro, F. Desempenho agrônômico e indicadores de sustentabilidade em tomate orgânico combinando diferentes práticas agroecológicas. *Nutrição. Cicl. Agroecosyst.* **2018**, *112*, 101–117. [\[CrossRef\]](#)
24. Shukla, MK; Lal, R.; Ebinger, M. Determinação de indicadores de qualidade do solo por análise fatorial. *Soil Tillage Res.* **2006**, *87*, 194–204. [\[Referência Cruzada\]](#)
25. Cardoso, EJBN; Vasconcellos, RLF; Bini, D.; Miyauchi, MYH; Santos, CA; Alves, PRL; Paula, AM; Nakatani, AS; Pereira, JM; Nogueira, MA Saúde do solo: Em busca de indicadores adequados. O que deve ser considerado para avaliar os efeitos do uso e manejo na saúde do solo? *Sci. Agric.* **2013**, *70*, 274–289. [\[CrossRef\]](#)
26. Cherubin, MR; Karlen, DL; Franco, AL; Cerri, CE; Tormena, CA; Cerri, CC Uma avaliação da expansão da cana-de-açúcar no Brasil, utilizando o Soil Management Assessment Framework (SMAF), sobre a qualidade do solo. *Soil Sci. Soc. Am. J.* **2016**, *80*, 215–226. [\[CrossRef\]](#)
27. Manici, LM; Castellini, M.; Caputo, F. Fungos que vivem no solo podem integrar indicadores físicos do solo em análises multivariadas de Agroecosistema mediterrâneo dominado por oliveiras antigas. *Ecol. Indic.* **2019**, *106*, 105490. [\[CrossRef\]](#)
28. Armenise, E.; Redmile-Gordon, MA; Stellacci, AM; Ciccicarese, A.; Rubino, P. Desenvolvimento de um índice de qualidade do solo para comparar a aptidão do solo para uso agrícola sob diferentes manejos no ambiente mediterrâneo. *Soil Till. Res.* **2013**, *130*, 91–98. [\[Referência Cruzada\]](#)
29. Andrews, SS; Karlen, DL; Cambardella, CA A estrutura de avaliação do manejo do solo: uma avaliação quantitativa da qualidade do solo método de avaliação. *Soil Sci. Soc. Am. J.* **2004**, *68*, 1945–1962. [\[CrossRef\]](#)
30. Rinot, O.; Levy, GJ; Steinberger, Y.; Svoray, T.; Eshel, G. Avaliação da saúde do solo: uma revisão crítica das metodologias atuais e uma nova abordagem proposta. *Sci. Total Environ.* **2019**, *648*, 1484–1491. [\[CrossRef\]](#)
31. Yemefack, M.; Jetten, VG; Rossiter, DG Desenvolvimento de um conjunto mínimo de dados para caracterizar a dinâmica do solo em sistemas de cultivo itinerante. *Soil Till. Res.* **2006**, *86*, 84–98. [\[CrossRef\]](#)
32. Mehmood, T.; Snipen, L.; Sæbø, S. Uma revisão de métodos de seleção de variáveis em regressão de mínimos quadrados parciais. *Química. Laboratório Intel. Sistemas.* **2012**, *118*, 62–69. [\[CrossRef\]](#)
33. Stellacci, AM; Castrignanò, A.; Troccoli, A.; Basso, B.; Buttafuoco, G. Seleção de bandas hiperespectrais ideais para discriminar o estado de nitrogênio em trigo duro: uma comparação de abordagens estatísticas. *Environ. Monit. Assess.* **2016**, *188*, 1–15. [\[CrossRef\]](#)
34. Pulido Moncada, M.; Gabriels, D.; Cornelis, WM Análise de indicadores de qualidade do solo baseada em dados usando dados limitados. *Geoderma* **2014**, *235*, 271–278. [\[Referência Cruzada\]](#)
35. De Paul Obade, V.; Lal, R. Um índice padronizado de qualidade do solo para diversas condições de campo. **2016**, *541*, 424–434. [\[Referência Cruzada\]](#) [\[PubMed\]](#)
36. Geladi, P.; Kowalski, BR Regressão por mínimos quadrados parciais: Um tutorial. *Anal. Química. Acta* **1986**, *185*, 1–17. [\[CrossRef\]](#)
37. Wold, S.; Sjöströma, M.; Lennart, E. Regressão PLS: Uma ferramenta básica da quimiometria. *Quimiometria. Laboratório Intel.* **2001**, *58*, 109–130. [\[Referência Cruzada\]](#)
38. Rencher, AC Métodos de Análise Multivariada, 2ª ed.; Universidade Brigham Young: Provo; John Wiley & Sons, Inc.: Nova York, NY, EUA, 2002.
39. Stellacci, AM; Castrignanò, A.; Diacono, M.; Troccoli, A.; Ciccicarese, A.; Armenise, E.; Gallo, A.; De Vita, P.; Lonigro, A.; Mastro, MA; et al. Abordagem combinada baseada em análise de componentes principais e análise discriminante canônica para investigar a resposta hiperespectral de plantas. **2012**, *7*, 247–253. [\[CrossRef\]](#)
40. UNESCO FAO. Mapa bioclimático da zona mediterrânea; (NS162/III, 22A); UNESCO: Paris, França; FAO: Roma, Itália, 1963; p. 60.
41. Ventrella, D.; Stellacci, AM; Castrignanò, A.; Charfeddine, M.; Castellini, M. Efeitos do manejo de resíduos culturais na produtividade do trigo duro de inverno em um experimento de longo prazo no sul da Itália. *Eur. J. Agron.* **2016**, *77*, 188–198. [\[CrossRef\]](#)
42. Equipe de Levantamento de Solos. Manual de Métodos de Campo e Laboratório para Levantamento de Solos; Relatório de Investigações de Levantamento de Solos nº 51, Versão 1.0; Burt, R., Ed.; Departamento de Agricultura dos EUA, Serviço de Conservação de Recursos Naturais: Washington, DC, EUA, 2009.
43. Vance, ED; Brooks, PC; Jenkinson, DS Um método de extração para medir a biomassa microbiana do solo C. *Soil Biol. Biochem.* **1987**, *19*, 703–707. [\[Referência Cruzada\]](#)
44. Burke, W.; Gabriels, D.; Bouma, J. Avaliação da Estrutura do Solo; Balkema: Roterdã, Holanda, 1986.

45. Dane, JH; Hopmans, JW 3.3. Retenção e Armazenamento de Água. Em Métodos de Análise de Solo, Métodos Físicos; Parte 4; Dane, JH, Topp, GC, Ed.; John Wiley & Sons: Nova York, NY, EUA, 2002; pp. 100-1 671–720.
46. Castellini, M.; Iovino, M. Funções de pedotransferência para estimar a curva de retenção de água em solos sicilianos. *Arq. Agron. Soil Sci.* **2019**, *65*, 1401–1416. [\[Referência Cruzada\]](#)
47. van Genuchten, MT Uma equação de forma fechada para prever a condutividade hidráulica de solos não saturados. *Soil Sci. Soc. Am. J.* **1980**, *44*, 892–898. [\[Referência Cruzada\]](#)
48. Wander, MM; Bollero, GA Avaliação da qualidade do solo em relação aos impactos da lavoura em Illinois. *Soil Sci. Soc. Am. J.* **1999**, *63*, 961–971. [\[CrossRef\]](#)
49. Sharma, KL; Mandal, Reino Unido; Srinivas, K.; Vittal, KPR; Mandal, B.; Grace, JK; Ramesh, V. Efeitos do manejo do solo a longo prazo na produtividade agrícola e na qualidade do solo em um Alfissolo de terras áridas. *Soil Tillage Res.* **2005**, *83*, 246–259. [\[CrossRef\]](#)
50. Thenkabail, PS; Enclona, EA; Ashton, MS; Van Der Meer, B. Avaliações de precisão do desempenho da banda de onda hiperespectral para aplicações de análise de vegetação. *Sensoriamento Remoto. Environ.* **2004**, *91*, 354–376. [\[CrossRef\]](#)
51. Lachenbruch, PA Análise Discriminante; Hafner Publishing: Nova York, NY, EUA, 1975.
52. Uddin, N. Previsão de dificuldades financeiras de clientes de crédito ao consumidor usando análise discriminante de dois grupos: um estudo de caso. *Int. J. Econ. Financ.* **2013**, *5*, 55–66. [\[CrossRef\]](#)
53. Uddin, N.; Meah, MS; Hossain, R. Análise discriminante como auxílio à seleção e rotatividade de recursos humanos Decisões de Minimização. *Int. J. Bus. Manag.* **2013**, *8*, 153–169. [\[CrossRef\]](#)
54. Software SAS/STAT; Versão 9.3; SAS Institute Inc.: Cary, NC, EUA, 2012.
55. Rosipal, R.; Kramer, N. Visão geral e avanços recentes em mínimos quadrados parciais. Subespaço, estrutura latente e seleção de recursos: Workshop de Perspectivas Estatísticas e de Otimização (SLSFS 2005); Artigos selecionados revisados (Anotações de aula em Ciência da Computação 3940); Springer: Berlim, Alemanha, 2006; pp. 34–51.
56. Colombo, C.; Palumbo, G.; Di Iorio, E.; Sellitto, VM; Comolli, R.; Stellacci, AM; Castrignanò, A. Variação do carbono orgânico do solo em paisagens alpinas (Norte da Itália) avaliada por espectroscopia de reflectância difusa. *Soil Sci. Soc. Am. J.* **2014**, *78*, 794–804. [\[Referência Cruzada\]](#)
57. Andersen, CM; Bro, R. Seleção de variáveis em regressão — Um tutorial. *J. Chemometr.* **2010**, *24*, 728–737. [\[CrossRef\]](#)
58. Chong, IG; Jun, CH Desempenho de alguns métodos de seleção de variáveis quando há multicolinearidade. *Chemometr. Intell. Lab. Syst.* **2005**, *78*, 102–112. [\[Referência Cruzada\]](#)
59. Wold, S. PLS para Modelagem Linear Multivariada, QSAR: Métodos Quimiométricos em Design Molecular. Em Métodos e Princípios em Química Medicinal; van de Waterbeemd, H., Ed.; Verlag-Chemie: Weinheim, Alemanha, 1994.
60. Castellini, M.; Stellacci, AM; Mastrangelo, M.; Caputo, F.; Manici, LM Estimativa das funções hidráulicas do solo em alguns olivais : implicações do manejo do solo para a economia de água em solos da Península de Salento (sul da Itália). *Agronomia* **2020**, *10*, 177. [\[Referência Cruzada\]](#)
61. Castellini, M.; Fornaro, F.; Garofalo, P.; Giglio, L.; Rinaldi, M.; Ventrella, D.; Vitti, C.; Vonella, AV Efeitos do plantio direto e do preparo convencional nas propriedades físicas e hidráulicas de solos de textura fina sob trigo de inverno. *Water* **2019**, *11*, 484. [\[CrossRef\]](#)
62. Castellini, M.; Giglio, L.; Modugno, F. Efeito do volume de solo amostrado na determinação da qualidade física do solo: Um estudo de caso em preparo convencional e plantio direto do solo sob trigo de inverno. *Soil Syst.* **2020**, *4*, 72. [\[CrossRef\]](#)
63. Paustian, K.; Six, J.; Elliott, ET; Hunt, HW Opções de gestão para reduzir as emissões de CO₂ de solos agrícolas. *Biogeochemistry* **2000**, *48*, 147–163. [\[CrossRef\]](#)
64. Diácono, M.; Persiani, A.; Testani, E.; Montemurro, F. Sustentabilidade das práticas agroecológicas na horticultura orgânica: Rendimento, uso de energia e pegada de carbono. *Agroecol. Sustain. Food Syst.* **2020**, *44*, 726–746. [\[CrossRef\]](#)
65. Laudicina, VA; Novara, A.; Gristina, L.; Badalucco, L. Dinâmica do carbono no solo afetada por sistemas de cultivo e preparos contrastantes de longo prazo em clima mediterrâneo semiárido. *Appl. Soil Ecol.* **2014**, *73*, 140–147. [\[CrossRef\]](#)
66. Bongiorno, G.; Bünemann, E.K.; Oguejiofor, CU.; Meier, J.; Gort, G.; Comans, R.; Mäder, P.; Brussaard, L.; de Goede, R. Sensibilidade de frações lábeis de carbono à lavoura e ao manejo da matéria orgânica e seu potencial como indicadores abrangentes da qualidade do solo em condições edafoclimáticas na Europa. **2019**, *99*, 38–50. [\[CrossRef\]](#)
67. López-Fando, C.; Pardo, MT Alterações nas características químicas do solo com diferentes práticas de cultivo em ambiente semiárido. *Lavra do solo Res.* **2009**, *104*, 278–284. [\[CrossRef\]](#)
68. Martin-Rueda, I.; Muñoz-Guerra, LM; Yunta, F.; Esteban, E.; Tenório, JL; Lucena, JJ Efeitos do preparo do solo e da rotação de culturas na produção de cevada e nutrientes do solo em um Haploxeralf Calcicortídico. *Cultivo do Solo Res.* **2007**, *92*, 1–9. [\[RefCruz\]](#)
69. Pavinato, PS; Dao, TH; Rosolem, CA Efeitos do manejo do solo e do fósforo no fósforo bioativo enzimaticamente lábil disponibilidade em Latossolos do Cerrado. *Geoderma* **2010**, *156*, 207–215. [\[CrossRef\]](#)
70. Balota, EL; Kanashiro, M.; Colozzi Filho, A.; Andrade, DS; Dick, RP Atividades enzimáticas do solo sob sistemas de preparo do solo e rotação de culturas de longo prazo em agroecossistemas subtropicais. *Braz. J. Microbiol.* **2004**, *35*, 300–306. [\[CrossRef\]](#)
71. Pieri, CJMG Fertilidade dos solos: um futuro para a agricultura na savana da África Ocidental; Springer: Berlim, Alemanha, 1992.
72. Vilýček, J.; Koco, Š. Índice integrado de qualidade do solo agrícola na Eslováquia. *J. Maps* **2018**, *14*, 68–76. [\[CrossRef\]](#)