



Uma API para a classificação de solos do Brasil

Glauber José Vaz¹, Luís de França da Silva Neto², Renato Nunes Lima³, Stanley Robson de Medeiros Oliveira¹

¹ Embrapa Informática Agropecuária, Campinas, SP, Brasil,
glauber.vaz@embrapa.br, stanley.oliveira@embrapa.br

² Embrapa Solos, Recife, PE, Brasil, luis.franca@embrapa.br

³ Unisal, Campinas, SP, Brasil, renato.nlma@gmail.com

RESUMO

O objetivo deste trabalho é disponibilizar por meio de uma API um sistema para a classificação de solos do Brasil, a fim de auxiliar no avanço da qualidade dos dados de solos e apoiar iniciativas de melhoria na gestão deste recurso. A API de classificação de solos facilita a curadoria de dados de solos e pode orientar um registro mais adequado destes dados. Com auxílio desta ferramenta, uma análise detalhada sobre a classificação de solos foi realizada em uma das mais importantes bases de dados de solos no país. Os resultados mostram que há uma necessidade muito grande de curadoria desses dados por parte dos especialistas do domínio e que a ferramenta considerada é poderosa aliada na melhoria da qualidade de dados de solos do Brasil.

PALAVRAS-CHAVE: SiBCS, Sistema especialista, Dados de solos, Curadoria de dados.

ABSTRACT

The objective of this work is to provide an API for the classification of soils of Brazil, in order to help improve the quality of soil data and support initiatives to improve the management of this resource. The soil classification API supports soil data curation and can guide a more appropriate record of these data. With this software, a detailed analysis on soil classification was performed in one of the most important soil databases in Brazil. The results show that there is a great need for the curation of this data by the domain specialists and that the tool considered is powerful in helping to improve the quality of soil data in Brazil.

KEYWORDS: SiBCS, Expert system, Soil data, Data curation.

INTRODUÇÃO

O Sistema Brasileiro de Classificação de Solos (SiBCS) é o sistema taxonômico oficial de classificação de solos do Brasil. Hierárquico, multicategórico e aberto, o sistema possibilita a classificação de todos os solos existentes no território nacional (DOS SANTOS et al., 2018). Em sua quinta edição, o SiBCS conta com 13 classes no primeiro nível categórico. Embora o sistema envolva seis níveis categóricos, o quinto e o sexto níveis não são obrigatórios.

Vaz et al. (2018) construíram um sistema especialista baseado nas regras do SiBCS para auxiliar profissionais que necessitem fazer a classificação de solos brasileiros. O sistema simula o raciocínio de um profissional especialista do domínio ao realizar a classificação de perfis de solo. Assim, pode ser utilizado para classificar perfis segundo o SiBCS em diferentes níveis categóricos, de acordo com os dados fornecidos, e também para validar perfis previamente classificados. Embora o SiBCS tenha sido publicado há muitos anos, não há outro programa de computador que faça essa classificação automática de solos brasileiros.

Para se obter maior alcance deste sistema especialista, ele foi disponibilizado via API (*Application Programming Interface*), uma forma de comunicação entre aplicações computacionais através de uma rede, predominantemente a Internet, e com o uso de uma linguagem comum (JACOBSON et al., 2011).

Por meio da AgroAPI, a Embrapa (2019) já vem disponibilizando algumas APIs para uso interno, negócios com organizações parceiras ou acesso aberto. Assim, promove a criação de valor na Agricultura Digital por meio de APIs, que possibilitam o compartilhamento de dados e serviços, facilitam a integração de sistemas de informação e o estabelecimento de acordos entre organizações e, ainda, viabilizam maior alcance dos resultados obtidos pela empresa e seus parceiros (VAZ et al., 2017). Embora com acesso restrito em um primeiro momento, a API criada neste trabalho também está disponível nesta plataforma de APIs da Embrapa.

Esta API pode auxiliar iniciativas que visem à melhoria da gestão do solo tanto nacionais quanto internacionais. O Programa Nacional de Solos do Brasil (Pronassolos), por exemplo, tem por principais objetivos a retomada da realização dos levantamentos pedológicos e o estabelecimento de uma base de dados integrada na qual as informações de solos estejam organizadas e sistematizadas para a consulta do público em geral (POLIDORO et al., 2016). Esta API pode auxiliar no Pronasolos facilitando a curadoria dos dados e indicando como os dados podem ser coletados a fim de facilitar seu processamento computacional. No contexto internacional, a Global Soil Partnership (GSP) foi estabelecida

para melhorar a governança e promover a gestão sustentável de solos (FOOD AND AGRICULTURAL ORGANIZATION OF THE UNITED NATIONS, 2019). Um de seus pilares de ação é avançar na qualidade e na quantidade de informações e dados de solos, mas, para isso, conta com a colaboração e o compartilhamento de informações locais. Assim, a API de classificação de solos também pode contribuir com esta e outras iniciativas globais.

A fim de se conduzir um teste preliminar no sistema, seus resultados de classificação foram comparados às classificações fornecidas pelo Banco de Dados de Informações Ambientais (BDiA). Este banco de dados é disponibilizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e reúne bases temáticas de recursos naturais do território brasileiro, dentre as quais a de Pedologia (IBGE, 2019).

O objetivo deste trabalho, portanto, é disponibilizar por meio de API um sistema especialista para a classificação de solos brasileiros desenvolvido com base no SiBCS, a fim de apoiar iniciativas de melhoria na gestão de solo, auxiliando, principalmente, no avanço da qualidade dos dados. Além disso, uma análise de consistência foi realizada com o auxílio desta ferramenta sobre os dados de uma das principais bases de solos do Brasil.

MATERIAL E MÉTODOS

O sistema especialista, desenvolvido na linguagem SWI Prolog, pode ser acessado via uma API disponível na AgroAPI, plataforma implementada com tecnologia para gerenciamento de APIs. Segundo Heffner (2018), este tipo de solução atende a, principalmente, três necessidades: (i) a otimização de valor por parte dos provedores de APIs, (ii) a gestão do relacionamento entre os provedores e os consumidores de APIs, e (iii) o estabelecimento de acordos sobre o uso e a segurança das APIs.

A comunicação com a API de classificação de solos é realizada com dados em JSON (*JavaScript Object Notation*), um formato leve para troca de dados que é fácil de ler e escrever para humanos e de interpretar e gerar por computadores. Neste formato, os dados são representados em texto e organizados em listas de valores e em coleções de pares formados por nome e valor (JSON.ORG, 2019).

Embora um arquivo JSON possa ser lido por um ser humano, sua edição direta pode não oferecer uma experiência boa para o usuário e ainda aumenta a probabilidade de erros nos dados. Assim, foi construída uma interface de usuário simples para a geração padronizada dos dados de entrada para a API. Por meio desta interface, é possível preencher os campos correspondentes aos dados de um perfil de solo e salvá-los em um arquivo JSON. Também é

possível carregar dados de um arquivo neste formato nos campos da interface para edição dos valores. Além de possibilitar a chamada ao sistema especialista de classificação de solos com os dados do formulário como entrada, esta interface faz alguns testes de consistência de valores para os campos. A ideia, no entanto, é que os clientes da API desenvolvam suas próprias interfaces para atendimento a necessidades específicas.

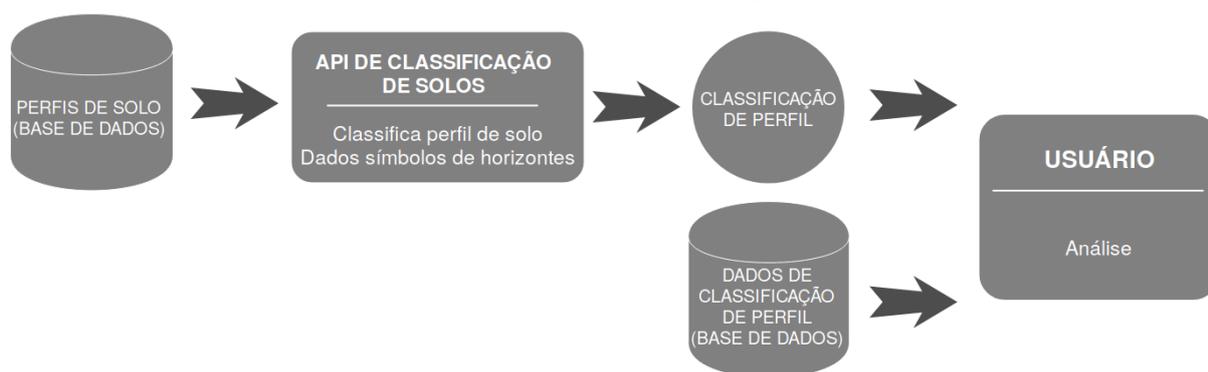
Os dados do BDiA foram utilizados para validar o sistema especialista. Eles são públicos, facilmente obtidos na Internet e contam com classificações para os perfis do primeiro ao quarto nível categórico do SiBCS. No entanto, devido à grande quantidade de perfis, uma análise mais detalhada foi realizada considerando-se apenas o primeiro nível. A fim de servir de entrada para o sistema especialista, os dados do BDiA foram convertidos para o formato JSON com várias alterações em seus campos e em seus conteúdos, de maneira a simplificar o processamento computacional e a padronizar os valores desses campos.

RESULTADOS E DISCUSSÃO

O sistema especialista disponibilizado via API assume que os valores dos atributos de solos são corretamente fornecidos, assim como as representações dos símbolos dos horizontes dos perfis de solo. Mesmo nos casos em que seria possível classificar o perfil independentemente dos símbolos de horizonte, estes são verificados e precisam estar corretamente especificados. Além disso, o sistema realiza a classificação de acordo com a quinta e última versão do SiBCS, podendo haver incompatibilidade com perfis que foram classificados com regras de edições anteriores.

A Figura 1 representa o processo de análise conduzido neste trabalho. Os resultados gerados pelo sistema disponibilizado via a API são comparados aos dados registrados. A Tabela 1 exhibe, para cada classe, a proporção de perfis cujas classificações obtidas via a API coincidem com as previamente anotadas no banco de dados utilizado.

Figura 1 – Análise de classificação de perfis de solo



Fonte: Adaptado de Vaz et al. (2018)

Tabela 1 – Resultados do sistema especialista com os dados do BDIA

CLASSE	TOTAL DE PERFIS	COINCIDE	NÃO COINCIDE	% COINCIDE
ORGANOSSOLOS	13	8	5	61.54 %
ESPODOSSOLO	65	45	20	69.23 %
PLANOSSOLOS	200	114	86	57 %
PLINTOSSOLO	247	226	21	91.5 %
VERTISSOLO	33	0	33	0 %
CHERNOSSOLO	86	23	63	26.74 %
GLEISSOLO	204	167	37	81.86 %
LUVISSOLO	76	51	25	67.11 %
NITOSSOLO	128	29	99	22.66 %
ARGISSOLO	1428	785	643	54.97 %
LATOSSOLO	1157	704	453	60.85 %
CAMBISSOLO	365	243	122	66.58 %
NEOSSOLO	465	437	28	93.98 %
TODOS	4467	2832	1635	63.4 %

Nota-se que em grande parte dos casos, há diferença na classificação do solo. Esse índice ainda é superior a uma análise prévia realizada, devido ao relaxamento de algumas restrições. Na verificação de planossolos, nitossolos e chernossolos, por exemplo, atributos de estrutura e de cor não foram verificados devido à ausência desses atributos em grande quantidade dos perfis. No teste de luvisolos, as regras referentes a planossolos foram desativadas, pois estrutura e cor são elementos importantes para diferenciar estas duas classes e, por isso, muitos dos perfis de luvisolos eram classificados como planossolos. Essa

supressão de algumas regras foi feita para verificar se as outras condições seriam atendidas nos casos em que os atributos mencionados fossem adequadamente anotados e satisfizessem às restrições.

Em outros casos, as regras foram mantidas mesmo quando tiveram grande impacto nos resultados. Por exemplo, o carbonato de cálcio equivalente é muito importante na verificação de chernossolos, mas a grande maioria dos perfis não trazem valores para este atributo. Ainda assim, as condições envolvendo-o foram mantidas. Se todos os perfis que dependem deste valor para ser classificado como chernossolo atendessem às restrições de carbonato de cálcio, as discrepâncias seriam reduzidas de 63 para 39. E se a verificação de planossolo também fosse desativada nestes mesmos casos, esse valor seria ainda menor, 29. Portanto, a anotação incorreta ou ausência de valores para um único atributo podem gerar grande impacto na classificação de solos.

A Tabela 1 também mostra a ordem com que as classes são testadas pelo sistema e, portanto, sua precedência. Assim, se determinado perfil atende às restrições de diferentes classes, o resultado corresponde à primeira delas. Um perfil sempre é classificado entre as treze classes consideradas, de maneira que, caso não atenda às condições de nenhuma das doze primeiras, automaticamente, é tratado como neossolo. Por exemplo, no conjunto de dados utilizado, há inúmeros perfis que não apresentam os dados de horizontes. Todos eles são classificados como neossolos.

As diferenças entre as classificações do banco de dados e do sistema especialista poderiam ser ocasionadas tanto pela implementação equivocada das regras do SiBCS no programa de computador quanto por inconsistências nos dados, como, por exemplo, valores incompatíveis para atributos específicos, simbologia de horizontes equivocados, ausência de valores para determinados atributos, não atendimento a restrições de determinada classe de solo e incompatibilidade de versões do SiBCS, entre outras.

Então, análises mais detalhadas foram realizadas para checar as causas dessas discrepâncias nas classificações. Nos casos em que foram identificados erros nas regras do software, estes foram corrigidos e os resultados esperados passaram a ser gerados. No entanto, em grande quantidade de perfis, o programa continuou a exibir classificações distintas das previamente fornecidas. A Tabela 2 apresenta os motivos porque cada perfil previamente anotado em determinada classe não foi classificado conforme a anotação.

Tabela 2 – Motivos pelos quais há diferença entre a classe anotada e a classe obtida pelo sistema especialista

CLASSE	MOTIVOS DE DIFERENÇA NA CLASSIFICAÇÃO
ORGANOSSOLOS	1 sem horizontes 1 sem horizonte H ou O 3 com quantidade de carbono orgânico insuficiente
ESPODOSSOLO	1 sem horizontes 19 sem horizonte Bh ou Bs
PLANOSSOLOS	14 sem horizontes 18 sem transição abrupta ou mudança textural abrupta 6 sem relação textural B/A de horizonte B plânico 22 sem horizonte textural Bt 14 sem horizonte A ou E antes de Bt 2 horizonte A ou E não está imediatamente antes de Bt 10 em classe precedente (plintossolos em casos com caráter plânico sem caráter sódico)
PLINTOSSOLO	5 sem horizontes 10 sem horizonte plântico, litoplântico ou concrecionário 6 horizonte anterior ao horizonte plântico, litoplântico ou concrecionário não atende
VERTISSOLO	33 sem horizonte com sufixo ‘v’
CHERNOSSOLO	5 sem horizontes 13 em classes precedentes (12 planossolo, 1 vertissolo) 13 não tem profundidade de A suficiente 3 não tem horizontes depois de A chernozêmico 6 não satisfazem restrições de saturação por bases 23 não tem Bi ou Bt, ou não atende a condições de rëndzico (sem dados de CaCO ₃)
GLEISSOLO	5 sem horizontes 10 em classe precedente (9 plintossolo, 1 chernossolo) 14 sem horizonte com sufixo ‘g’ 5 horizonte anterior ao glei não atende 3 sem profundidade suficiente de horizonte glei
LUVISSOLO	9 em classe precedente (plintossolos) 7 não atende a condições de saturação por bases alta ou argila de atividade alta 8 sem horizontes Bt 1 sem profundidade suficiente de Bt
NITOSSOLO	2 sem horizontes 10 em classes precedentes (6 planossolo, 4 luvisolo) 50 sem horizontes Bt 9 sem profundidade suficiente de Bt 1 sem textura argilosa (sem dados de argila) 23 sem registro de cerosidade 4 com cerosidade que não representa nitossolo
ARGISSOLO	70 sem horizontes 436 em classes precedentes (255 plano, 127 plinto, 44 nito, 6 cherno, 2 luvi, 2 gleissolo) 135 não tem Bt imediatamente abaixo de A ou E 2 não apresenta argila baixa ou saturação por bases baixa ou caráter aluminico
LATOSSOLO	87 sem horizontes 44 em classes precedentes (1 plano, 25 plinto, 16 argi, 2 nito) 322 não tem Bw abaixo de A
CAMBISSOLO	37 sem horizontes 28 em classe precedente (1 gleissolo, 21 plintossolo, 6 argissolo) 57 sem horizonte Bi imediatamente abaixo de A, E, O ou H
NEOSSOLO	28 em classes precedentes (2 plano, 11 plinto, 5 glei, 2 nito, 2 argi, 5 cambi, 1 latossolo)

Para as classes listadas de organossolos a nitossolos, a análise foi bastante minuciosa, de maneira que os motivos das discordâncias nas classificações foram até mais detalhadas. A partir de argissolos até neossolos, devido ao maior número de perfis, os motivos levantados são um pouco mais genéricos e contaram com um maior grau de automação. Porém, em todos os casos, houve um trabalho humano na análise desses dados. Portanto, embora sujeita a erros, essa análise fornece um retrato geral da consistência dos dados utilizados e de suas classificações.

Observa-se uma grande quantidade de perfis que não possuem dados relacionados a seus horizontes, o que caracteriza falha no registro dessas informações. Nos perfis classificados como neossolos, ocorrências desse tipo também foram verificadas, mas são considerados neossolos porque não atendem às restrições de nenhuma das outras classes. Diferenças nas classificações devido à simbologia de horizontes também foram bastante comuns neste conjunto de dados. Em trabalhos futuros, isso pode ser amenizado porque os símbolos também podem ser inferidos a partir dos dados coletados no campo. Também foram comuns situações em que os perfis assumiram classes testadas previamente pelo programa, quando também atenderiam às restrições da classe anotada. A desativação de algumas regras importantes na distinção das classes explica a maior parte desse efeito, porém, esta medida somente foi tomada porque valores de atributos essenciais não foram registrados para uma grande quantidade de perfis, com destaque para os de cor e de estrutura. A Tabela 2 ainda exibe uma série de informações com riqueza de detalhes e que possibilitam a compreensão das inconsistências dos dados de solos registrados.

Este trabalho de análise dos dados contou com auxílio de pedólogos, mas foi realizado predominantemente por especialistas de computação e análise de dados, leigos no domínio de solos. Portanto, tanto o software quanto esta análise demandam uma validação mais abrangente por parte dos especialistas em classificação de solos.

CONCLUSÕES

Os resultados mostrados neste trabalho fornecem um quadro geral da qualidade dos dados considerados. A base utilizada é uma das mais importantes, mais utilizadas e de maior qualidade para se obter dados de perfis de solos do Brasil e suas classificações. No entanto, ainda há uma necessidade muito grande de curadoria desses dados por parte dos especialistas desse domínio. Também, é importante que novos levantamentos de solos sejam realizados de maneira que os dados obtidos sejam registrados de uma maneira mais padronizada e mais

adequada para o processamento computacional, reduzindo-se assim esses erros tão presentes nas bases de dados disponíveis atualmente.

A API de classificação de solos considerada neste trabalho pode auxiliar na curadoria dos dados existentes e na verificação da consistência de novos dados coletados. O padrão de entrada necessário para o uso correto desta API também pode nortear o registro dos dados de forma padronizada e de maneira a abranger todos os dados necessários para a correta classificação do solo. Portanto, embora ainda demande uma maior validação por parte dos especialistas em classificação de solos, trata-se de uma poderosa ferramenta para a melhoria da qualidade de dados de solos no Brasil.

AGRADECIMENTOS

Esta pesquisa foi apoiada pela EMBRAPA. Projeto SmartSolos [SEG 23.14.01.017.00.00].

REFERÊNCIAS

- DOS SANTOS, H. G. et al. Sistema brasileiro de classificação de solos. Brasília, DF: Embrapa, 2018.
- EMBRAPA. AgroAPI: Plataforma de APIs da Embrapa. Disponível em: <<https://www.agroapi.cnptia.embrapa.br>>. Acesso em: 27 mai. 2019.
- FOOD AND AGRICULTURAL ORGANIZATION OF THE UNITED NATIONS. Global Soil Partnership. Disponível em: <www.fao.org/global-soil-partnership>. Acesso em: 18 jun. 2019.
- HEFFNER, R. (2018). The Forrester Wave TM: API Management Solutions, Q4 2018. Forrester Research, Inc, 2018.
- IBGE. BDIA: Banco de dados de informações ambientais. Disponível em: <<https://bdiaweb.ibge.gov.br>>. Acesso em 31 mai. 2019.
- JACOBSON, D.; WOODS, D.; BRAIL, G. APIs: A strategy guide. O'Reilly Media, Inc., 2011.
- JSON.ORG. Introducing JSON. Disponível em: <<http://json.org/>>. Acesso em 28 jun. 2019.
- POLIDORO, J. C. et al. Programa Nacional de Solos do Brasil (PronaSolos). Rio de Janeiro: Embrapa Solos, 2016.
- VAZ, G. J.; APOLINÁRIO, D. R. de F.; CORREA, J. L.; VACARI, I.; GONZALES, L. E.; DRUCKER, D. P.; BARIANI, J. M.; EVANGELISTA, S. R. M.; ROMANI, L. A. S. AgroAPI: criação de valor para a Agricultura Digital por meio de APIs. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 11., 2017, Campinas. Ciência de dados na era da agricultura digital: anais. Campinas: Editora da Unicamp: Embrapa Informática Agropecuária, 2017. p. 59-68. SBIAgro 2017.

VAZ, G. J.; SILVA NETO, L. de F. da; OLIVEIRA, S. R. de M.; BOTELHO, F. P.; ARAUJO FILHO, J. C. de. Development of an expert system for classification of Brazilian soil profiles. In: WORLD CONGRESS OF SOIL SCIENCE, 21., 2018, Rio de Janeiro. Soil science: beyond food and fuel: abstracts. Viçosa, MG: SBCS, 2018. Não paginado. WCSS 2018.