BANCO DE DADOS DISTRIBUÍDO NA GESTÃO DE DADOS AGROMETEOROLÓGICOS

Mateus Felipe Ribeiro – Universidade Estadual de Ponta Grossa – UEPG – 240401400006@uepg.br
Wesley Vinicius Fernandes –Universidade Estadual de Ponta Grossa – UEPG – 240401400010@uepg.br
Maria Beatriz Macedo Simon Sola – Universidade Estadual de Ponta Grossa – UEPG – 24040140005@uepg.br
Maria Salete Marcon Gomes Vaz – Universidade Estadual de Ponta Grossa – UEPG – salete@uepg.br
Gabriel Passos de Jesus – Universidade Estadual de Ponta Grossa – UEPG – gabrielpasje@gmail.com

Resumo: Os bancos de dados armazenam e utilizam dados na gestão de informações agrometeorológicas para tomada de decisões agrícolas. A agricultura depende da predição climática para determinar culturas e otimizar práticas, como irrigação e manejo de pragas, visando sustentabilidade e eficiência. Dados meteorológicos são coletados por estações automáticas, sensores orbitais e radares, processados e armazenados em bancos de dados distribuídos para garantir integridade e consistência. A implementação de banco de dados distribuído facilita a gestão eficiente dos dados agrometeorológicos, permitindo sua coleta, processamento e integração de maneira eficaz. Este trabalho explora o papel desses bancos de dados, avaliando a coleta, o processamento e a integração de dados, além de identificar as vantagens e desafios técnicos da implementação. Pesquisas recentes aplicam técnicas que combinam o uso de sensores com bases distribuídas para melhoria da gestão da irrigação e aumento da produtividade das plantações, evidenciando ganhos na produção. Algoritmos de aprendizado de máquina com redes neurais são utilizados para prever a produção agrícola, apoiados por grandes bases de dados distribuídas. Assim, os dados agrometeorológicos e os bancos de dados distribuídos desempenham um papel importante na modernização agrícola, impulsionando práticas sustentáveis e tomada de decisões para atingir a eficiência agrícola.

Palavras-chave: Bancos de dados distribuídos. Gestão. Dados meteorológicos. Eficiência. Agricultura.

DISTRIBUTED DATABASE FOR AGROMETEOROLOGICAL DATA MANAGEMENT

Abstract: *Databases are essential for storing and using information, especially in the management of agrometeorological data, which is crucial for agricultural decision-making. Agriculture depends on climate prediction to determine ideal crops and optimize practices such as irrigation and pest management, aiming for sustainability and efficiency. Weather data is collected by automatic stations, orbital sensors and radars, processed and stored in distributed databases (DBD) to ensure integrity and consistency. The implementation of DDBs facilitates the efficient management of agro-meteorological data, allowing it to be collected, processed and integrated effectively. This work aims to explore the role of DDBs in the management of agrometeorological data, evaluating the collection, processing and integration of this data, as well as identifying the advantages and technical challenges of implementing such databases. Recent studies have applied techniques combining the use of sensors with BBD to improve irrigation management and increase crop productivity, showing significant gains in production. In addition, machine learning algorithms, such as neural networks, are used to predict agricultural production, supported by large distributed databases. Thus, agrometeorological data and BDDs play a crucial role in the modernization of agriculture, driving sustainable practices and informed decisions, by combining this information we have better agricultural efficiency.*

Key-words: *Distributed databases. Management. Meteorological data. Efficiency. Agriculture.*

1. Introdução

Os bancos de dados são ferramentas para armazenar dados e são utilizados para solução para um determinado problema, como na gestão de dados meteorológicos, na tomada de decisão no setor agrícola (PLACIDA; OGHENEKARO, 2021).

A agricultura é uma das principais utilizadoras dos dados meteorológicos, pois com informações da predição climática é possível determinar com precisão qual cultura para determinada região e época do ano. Esses dados, segundo Raj *et al*. (2021), são eficazes para o agricultor manejar pragas, sem desperdiçar insumos, e realizar irrigação adequada, por exemplo, visando sempre a economia e a sustentabilidade.

Para isso, os dados meteorológicos são gerados por meio de estações meteorológicas automáticas ou convencionais, e por meio de sensores orbitais ou radares. Após coletados, eles são enviados para os chamados distritos, uma filial regional do INMET (Instituto Nacional de Meteorologia), onde são processados e encaminhados para a sede do INMET, em Brasília - DF, para serem armazenados em um banco de dados oficial. Em conjunto com outros indicadores como solo, cultura a ser implantada, pragas, doenças, entre outros, formam os dados agrometeorológicos (FIGUEIREDO, 2019).

Para armazenamento desses dados, o uso de bancos de dados distribuídos são adequados para manter a integridade e consistência na gestão dos dados agrometeorológicos. Estes banco de dados estão distribuídos, não estando em um único servidor, espalhados por vários servidores, chamados nós, fisicamente no mesmo local ou distribuído por uma rede de computadores (GUO; JIANG; ZHAO, 2023).

Desse modo, este artigo tem como objetivo explorar o papel dos bancos de dados distribuídos na gestão de dados agrometeorológicos, visto que entender como essa classificação de base de dados atua na gestão dos dados agrometeorológicos é importante na tomada de decisões agrícolas.

2. Material e Métodos

Este trabalho consistiu em uma revisão da literatura acerca do uso dos bancos de dados distribuídos para a gestão de dados agrometeorológicos, com o intuito de compreender como a literatura se apresenta no período de 2014 a 2024.

A pesquisa foi conduzida através do ScienceDirect e do Portal de Periódicos da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), base de dados referência no país e que possui estudos nacionais e internacionais, utilizando termos de pesquisa como "*distributed databases*", "*agrometeorological data*", "*data management*" e "*agriculture"*.

O passo seguinte consistiu em extrair os dados dos materiais selecionados, realizando uma ficha de leitura contendo os seguintes tópicos: objetivo do estudo, hipóteses e questões colocadas pelo(s) autor(es), metodologia empregada, principais resultados, discussão realizada com os resultados de trabalhos correlatos, e as conclusões do estudo.

Por fim, os dados extraídos foram agrupados em três temas: Banco de Dados Distribuídos; Gestão de Dados Agrometeorológicos; e Banco de Dados Agrometeorológicos Distribuídos. Tais temas sintetizam o estado e os conceitos básicos da literatura revisada.

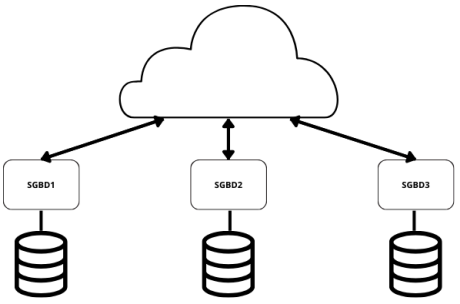
3. Banco de Dados Distribuído

Banco de dados distribuído é entendido por Guo, Jiang e Zhao (2023) como aquele banco de dados localizado em diferentes máquinas ou locais, mas que aparecem como um banco de dados centralizado para o usuário final.

A informação é compartilhada por um conjunto de computadores, com vários servidores trabalhando em sincronia para atender às necessidades de múltiplos usuários, em vez de ser gerenciada por uma única base de dados centralizada.

Essas máquinas estão interligadas em um sistema distribuído por conexões sem fio ou outros meios de comunicação que transmitem dados em alta velocidade. A Figura 1 mostra um exemplo de comunicação realizado por um banco de dados distribuído, onde o usuário se comunica com a "nuvem" e em seguida é realizada uma solicitação da informação para o primeiro SGBD disponível.

Figura 01 – Comunicação do BDD com a rede (usuário)



Fonte: os autores, 2024

Além disso, Placida e Oghenekaro (2021) complementam que, ao contrário dos bancos de dados centralizados, onde todos os dados são armazenados em um único local e

manipulados sequencialmente em uma única transação, os bancos de dados distribuídos processam solicitações por meio de arquiteturas cliente/servidor. Esse tipo de sistema é comumente categorizado como banco de dados distribuído homogêneo ou heterogêneo.

Desse modo, Ali e Nugraha (2019) ressaltam que os banco de dados distribuídos oferecem vantagens ao melhorarem a flexibilidade no desenvolvimento de aplicações, pois a transparência na distribuição de dados facilita o desenvolvimento e a manutenção em locais geograficamente dispersos. Mostram, ainda, que a confiabilidade e a disponibilidade são aumentadas, visto que as falhas são isoladas ao local de origem sem afetar toda a rede, permitindo que outros locais continuem operando, e a replicação de dados em vários locais evita interrupções totais do sistema.

Os autores apontam, ainda, que o desempenho é aprimorado, uma vez que os dados são mantidos próximos de onde são necessários, reduzindo a concorrência por recursos e permitindo a execução paralela de consultas, o que resulta em menor espera e maior eficiência geral.

Apesar das vantagens, os bancos de dados distribuídos apresentam algumas desvantagens, principalmente em relação à complexidade, maior em comparação com sistemas centralizados, já que gerenciar a duplicação de dados e garantir a integridade em uma rede distribuída pode levar a maiores custos de aquisição e manutenção devido à necessidade de hardware adicional e conectividade contínua. Além disso, a segurança é um desafio, considerando que o acesso a dados replicados deve ser gerenciado em múltiplos locais, e a rede em si precisa ser protegida contra vulnerabilidades (ALI; NUGRAHA, 2019).

Todos esses fatores apontados por Ali e Nugraha (2019) mostram que a implementação e a gestão de SGBDs distribuídos são mais complexas e onerosas.

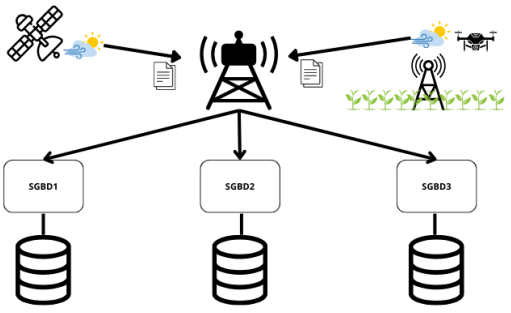
4. Gestão de Dados Agrometeorológicos

Dado meteorológico corresponde a todo dado que faz referência a um estado do clima de uma região, sendo esse dado um valor numérico ou textual. Os valores numéricos podem ser temperatura, umidade, velocidade do vento e índice UV (raios ultravioleta), já os valores textuais dizem respeito a descrição da sensação térmica de forma mais humanizada, ou seja, informa se o clima está quente ou frio, por exemplo (BLOH *et al*., 2023; MIA *et al*, 2023; LINGWAL *et al*., 2024).

Diante disso, o uso de ferramentas de coleta de dados como ferramentas de internet das coisas, veículos aéreos não tripulados, e de dados provindos de satélites são fundamentais para que se tenha acesso a essas informações valiosas (SAIZ-RUBIO e ROVIRA-MÁS, 2019; ADAMIDES *et al.*, 2020; IDOJE *et al.*, 2021; MOYSIADIS *et al.*, 2021; MOHAMED *et al.*, 2021; WANG *et al.*, 2021; DHANARAJU *et al.*, 2022; ALWIS *et al.*, 2022; JOSHI *et al.*, 2023).

Por sua vez, a coleta de dados meteorológicos provindos de satélites ou de estações fixas é constante. Diferente de estações meteorológicas móveis, coletas constantes demandam uma grande capacidade de armazenamento. Na Figura 2, pode-se visualizar um exemplo de comunicação, onde os sensores e dispositivos de internet das coisas enviam as informações para uma estação principal e em seguida essas informações são armazenadas no banco de dados distribuído, ou seja, cada instância de SGBD irá receber a mesma informação da estação, promovendo assim a confiabilidade e estabilidade dos dados.

Figura 2 – Comunicação dos sensores e dispositivos Internet das Coisas com Bancos de Dados Distribuídos



Fonte: os autores, 2024

Estes dados são distribuídos entre redes de comunicação ou até mesmo estações e servidores, até serem armazenados em banco de dados convencionais ou distribuídos, tudo

isso através dos dispositivos da intenet das coisas. A velocidade de transmissão entre dispositivos tem evoluído e confiável, tornando-se importante para conduzir os dados até seu armazenamento (IDOJE *et al.*, 2021).

Os sensores realizam leituras de diversos comprimentos de ondas, como por exemplo, sensores infravermelho, imagens e também geolocalização. Diversos trabalhos utilizam dados combinados de imagens (SONG *et al.*, 2022) e dados meteorológicos para, por exemplo, estimarem a produção de uma cultura de soja, como estudado por Bloh *et al.* (2023). Esta combinação forma os chamados dados agrometeorológicos.

No setor agrícola há um interesse crescente em gerar e usar dados combinados com dados meteorológicos, em especial na tomada de decisão. Segundo Chamorro-Padial; García e Gil (2024), a coleta de dados agrícolas pode incluir informações sobre clima, solo, rendimento das colheitas e práticas de manejo, essenciais para tomar decisões, visando reduzir custos, aumentar a produtividade e mitigar os impactos ambientais.

A coleta sistemática de dados permite aos agricultores e pesquisadores identificar padrões, prever tendências e desenvolver estratégias eficazes para o manejo agrícola.

5. Banco de Dados Agrometeorológicos Distribuídos

Os bancos de dados distribuídos têm a capacidade de processar e analisar grandes conjuntos de dados provenientes de diversas fontes, como no estudo realizado por Culman *et al.*, (2019). Os autores utilizaram uma rede de sensores para coletar dados agrometeorológicos em plantações de palma, distribuindo os sensores espacialmente no campo para medir continuamente dados agrometeorológicos, como umidade do solo, déficit de pressão de vapor, temperatura do ar, radiação solar e precipitação. Esses dados são transmitidos e armazenados em um banco de dados distribuído.

A metodologia usada envolveu a integração desses dados utilizando técnicas de *Data Fusion*, que combinam dados de múltiplas fontes para obter uma estimativa precisa e confiável, através do método de Inferência de *Dempster-Shafer*, que permite a combinação de informações de diferentes sensores, considerando diferentes níveis de incerteza. Essa análise possibilita inferir o estado hídrico das plantas e do solo, permitindo uma tomada de decisão precisa sobre a irrigação.

Os resultados das simulações mostraram um aumento de 27% na produção de cachos de frutos frescos nas áreas onde a irrigação foi gerida pelo método de inferência em comparação com áreas sem irrigação, demonstrando a eficácia da metodologia em melhorar a gestão da irrigação e a produtividade das plantações.

Já no estudo realizado por Bharambe *et al.* (2023), houve coleta de dados agrometeorológicos e geoespaciais, como precipitação, temperatura e dados topográficos, de fontes como o *Distributed Active Archive Center* (*DAAC*), centro responsável pelo armazenamento e disseminação de dados científicos coletados por várias missões da NASA.

A forma proposta neste estudo envolve a análise de índices padronizados de seca e projeções climáticas para avaliar o risco de seca em sub-bacias específicas do rio Godavari, na Índia. Os resultados mostraram um aumento na severidade e frequência das secas futuras, com 75% da área da bacia, apresentando maior vulnerabilidade à seca, resultando em perdas significativas na produção de arroz. Isso possibilitou uma avaliação precisa dos impactos das mudanças climáticas em diferentes sub-bacias do rio Godavari, identificando áreas de maior vulnerabilidade e projetando as condições futuras das secas.

Uma outra aplicação para os dados agrometeorológicos é na predição de produção de grãos, tendo como auxílio algoritmos de aprendizado de máquinas, conforme Bloh *et al.* (2023) e Mia *et al.* (2023) mostraram ao fazer uso da aplicação de redes neurais artificiais nesses dados. Os trabalhos ressaltaram que o agricultor se beneficiará das informações, como estimativa para a produção da sua cultura, estimando financeiramente a renda da colheita na estação.

6. Conclusões

A implementação de bancos de dados distribuídos e a aplicação de técnicas avançadas de análise de dados transformam a agricultura, tornando-a resiliente, produtiva e sustentável. Ao aproveitar ao máximo os dados agrometeorológicos, os agricultores podem otimizar suas operações, reduzir desperdícios e melhorar a qualidade e a quantidade de suas colheitas, contribuindo para a segurança alimentar e o desenvolvimento sustentável do setor agrícola.

Com isso, conclui-se que os bancos de dados distribuídos são um componente para a agricultura, permitindo a integração e análise de grandes volumes de dados agrometeorológicos, melhorando a tomada de decisões, aumentando a eficiência operacional e promovendo a sustentabilidade.

Ao investir em tecnologias de dados avançadas, o setor agrícola pode enfrentar os desafios com maior confiança e eficácia, garantindo a produção de alimentos suficiente para a população global em crescimento e preservando os recursos naturais para as gerações futuras.

Referências

ALI, A. H.; NUGRAHA, R. A. **Application of distributed databases for information systems fertilizer management.** IOP Conference Series: Materials Science and Engineering, v. 662, p. 022045, 20 nov. 2019.

ALWIS, S. et al. **A survey on smart farming data, applications and techniques.** Computers in Industry, v. 138, p. 103624, 2022.

BHARAMBE, K. P. et al. I**mpacts of climate change on drought and its consequences on the agricultural crop under worst-case scenario over the Godavari River Basin, India.** Climate Services, v. 32, p. 100415, 1 dez. 2023.

BLOH, M. et al. **Machine learning for soybean yield forecasting in Braz**il. Agricultural and Forest Meteorology, v. 341, p. 109670, 2023.

CHAMORRO-PADIAL, J.; GARCÍA, R.; GIL, R. **A systematic review of open data in agriculture.** Computers and Electronics in Agriculture, v. 219, p. 108775, 1 abr. 2024.

CULMAN, M. et al. **Using agrometeorological data to assist irrigation management in oil palm crops: A decision support method and results from crop model simulation**. Agricultural Water Management, v. 213, p. 1047–1062, 1 mar. 2019.

DHANARAJU, M. et al. **Smart farming: Internet of things (IoT)-based sustainable agriculture**. Agriculture, v. 12, p. 1745, 2022.

FIGUEIREDO, J. L. A. **Agrometeorologia.** Londrina: Editora e Distribuidora Educacional S.A., 2019. 168 p.

GUO, Z.; JIANG, G.; ZHAO, J. **Data processing method of distributed parallel database system based on wireless network.** Wireless Communications and Mobile Computing, v. 2023, p. 1–1, 20 jan. 2023.

JOSHI, A. et al. **Remote-sensing data and deep-learning techniques in crop mapping and yield prediction: A systematic review**. Remote Sensing, v. 15, p. 2014, 2023.

LINGWAL, S.; BHATIA, K. K.; SINGH, M. **A novel machine learning approach for rice yield estimation.** Journal of Experimental & Theoretical Artificial Intelligence, v. 36, n. 3, p. 337-356, 2024.

MIA, M. S. et al. **Multimodal deep learning for rice yield prediction using UAV-based multispectral imagery and weather data.** Remote Sensing, v. 15, p. 2511, 2023.

PLACIDA O., O.; OGHENEKARO A., P. **Distributed database management system (DBMS) architectures and distributed data independence**. International Journal of Computer Science and Mobile Computing, v. 10, n. 1, p. 23–48, 28 jan. 2021.

RAJ, M. et al. **A survey on the role of Internet of Things for adopting and promoting Agriculture 4.0.** Journal of Network and Computer Applications, v. 187, p. 103107, ago. 2021.

SONG, X. et al. **Annual 30 m soybean yield mapping in Brazil using long-term satellite observations, climate data and machine learning**. Agricultural and Forest Meteorology, v. 326, p. 109186, 2022.