

Desenvolvimento de um Sistema de Irrigação Inteligente: cultivando alface americana com o auxílio de impressora 3D e aprendizado por reforço

Hudson Ferreira Luiz¹

¹ ¹ Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais (IFMG)
Sabará – MG – Brasil

hudsonferreira2501@gmail.com

Abstract. *This work presents a proposal for a smart irrigation system for water management in vegetable gardens, integrating soil moisture sensors, evapotranspiration models, and artificial intelligence techniques. The system utilizes an ESP32 microcontroller for data collection, integrated with a backend server application responsible for data processing and storage, which communicates with an artificial intelligence model for decision-making.*

The evapotranspiration forecast is estimated using the Penman-Monteith model and meteorological data. Soil moisture is obtained from sensors calibrated based on agronomic parameters, such as field capacity (FC) and permanent wilting point (PWP).

The training of the decision agent is carried out in a simulated environment using a Beam Search heuristic, which allows exploring multiple simulation sequences and selecting those with the highest cumulative reward. During this training, the agent learns to adjust the irrigation volume based on environmental conditions and soil moisture.

Following this training, the learned policy is applied deterministically, correcting the irrigation volume proposed by the evapotranspiration forecast to maintain soil moisture close to the target level.

The results highlight a clear trade-off between the evaluated training approaches. Agent training guided by the Beam Search local heuristic algorithm showed faster convergence with a stable irrigation policy safe from water risks. On the other hand, training via a Softmax policy, although requiring a higher computational cost and a prolonged training time, shows increasing potential and a higher probability of finding the global optimum.

Resumo. *Este trabalho apresenta a proposta de um sistema de irrigação inteligente para manejo hídrico em hortas, integrando sensores de umidade do solo, modelos de evapotranspiração e técnicas de inteligência artificial. O sistema utiliza um microcontrolador ESP32 para a coleta de dados, integrado a uma API REST, responsável pelo processamento e armazenamento das informações, a qual se comunica com um modelo de inteligência artificial para a tomada de decisões. A previsão de evapotranspiração é estimada a partir do modelo de Penman-Monteith utilizando dados meteorológicos. A umidade do solo é obtida a partir de sensores calibrados com base em parâmetros agrônômicos,*

como capacidade de campo (FC) e ponto de murcha permanente (PWP). O treinamento do agente de decisão é realizado em ambiente simulado utilizando a busca heurística Beam Search, que permite explorar múltiplas sequências de simulação e selecionar aquelas com maior recompensa acumulada. Durante este treinamento, o agente aprende a ajustar o volume de irrigação com base nas condições ambientais e na umidade do solo. Após este treinamento, a política aprendida é aplicada de forma determinística, corrigindo o volume de irrigação proposto pela previsão de evapotranspiração, para manter a umidade do solo próxima ao nível alvo. Os resultados evidenciam um claro trade-off entre as abordagens de treinamento avaliadas. O treinamento do agente guiado pelo algoritmo de busca heurística local beam-search mostrou uma convergência mais rápida com uma política de irrigação estável e segura contra riscos hídricos. Por outro lado, o treinamento via política Softmax, embora exija um custo computacional maior, e um tempo de treinamento prolongado, mostra potencial crescente e probabilidade maior para encontrar o ótimo global.

1. Introdução

A gestão da irrigação impacta diretamente a produção agrícola, podendo resultar tanto em desperdício de água quanto em déficit hídrico, o que compromete o desenvolvimento das plantas [Da Silva et al. 2016]. Nesse contexto, o uso de inteligência artificial, em particular modelos de aprendizado por reforço, se destaca como uma abordagem capaz de lidar com a complexidade do sistema, permitindo estimar a quantidade ideal de irrigação a partir do aprendizado com dados históricos do ambiente. Contudo, a escassez desse tipo de dado ainda é um empecilho para a implantação desse tipo de solução [Liu et al. 2025].

Além de abordagens baseadas em dados, modelos físico-matemáticos também desempenham um papel importante na estimativa da necessidade hídrica. O modelo de Penman-Monteith, recomendado pela FAO (Organização das Nações Unidas para a Alimentação e a Agricultura) [Allen et al. 1998], é um dos principais guias sobre o cálculo de demanda hídrica das culturas. Ele permite estimar a evapotranspiração a partir de variáveis climáticas, fornecendo uma base teórica para o cálculo da demanda de água das culturas.

Diante desse cenário, surge a questão: como manter o volume de água no solo próximo do nível ideal para a cultura, considerando que os modelos de previsão de demanda hídrica dependem de dados meteorológicos sujeitos a imprecisão, e sem depender de grandes volumes de dados históricos de campo, escassos neste domínio? Este trabalho busca responder a essa questão por meio da combinação de um modelo físico-matemático de estimativa de demanda hídrica (Penman-Monteith) com um agente de aprendizado por reforço, treinado em ambiente simulado, capaz de corrigir o volume de irrigação proposto de acordo com as condições reais observadas no canteiro.

Usamos um microcontrolador ESP32 para possibilitar a coleta contínua de dados do solo por meio de sensores, viabilizando a formação de uma base de dados para análise e treinamento de modelos de aprendizado por reforço, capazes de aprender uma política de decisão para irrigação mediante variáveis climáticas, do solo e do tipo de cultura.

Um dos principais desafios na aplicação do aprendizado por reforço é o tempo

elevado de treinamento do agente. Além disso, a ocorrência frequente de falhas durante o processo de aprendizado — característica inerente a essa classe de modelos — pode resultar na perda significativa de recursos no contexto agrícola. Para contornar essa limitação, pode-se recorrer a simulações em ambiente virtual, abordagem adotada neste trabalho.

Neste trabalho, o agente foi treinado a partir de simulações no software *AquaCrop*, alimentado com dados meteorológicos históricos obtidos da plataforma NASA POWER (Prediction Of Worldwide Energy Resources), acessados via Data Access Viewer. Para otimizar as simulações e avaliar diferentes estratégias de tomada de decisão, utilizou-se o algoritmo heurístico de busca *Beam Search*, cujos resultados foram comparados aos obtidos pela política probabilística baseada na função *Softmax*.

O *AquaCrop*, desenvolvido pela Organização das Nações Unidas para a Alimentação e a Agricultura (FAO) [Salman et al. 2025], é um software gratuito concebido para simular o crescimento e a resposta do rendimento das culturas em função do uso da água. O modelo monitora dinamicamente diversas variáveis biofísicas do sistema solo-planta-atmosfera sob diferentes cenários de manejo hídrico. Neste trabalho, embora o simulador forneça múltiplos parâmetros de saída, utilizou-se exclusivamente o volume de água contido no solo para compor o estado de treinamento do agente. Essa escolha justifica-se pela aderência ao cenário real de aplicação, visto que o teor de umidade do solo corresponde à única variável física coletada diretamente no canteiro por meio dos sensores capacitivos instalados.

O *Beam Search* é um algoritmo de busca heurística em grafos e árvores de decisão que expande os estados mais promissores a partir de uma largura de busca fixa (k), denominada *beam width* [Russell and Norvig 2020, Lowerre 1976]. Em problemas de decisão sequencial, a exploração exaustiva de todas as combinações de irrigação ao longo do ciclo da cultura levaria a uma explosão combinatória impraticável. Ao limitar a busca às k melhores trajetórias a cada passo temporal ($k = 3$), o algoritmo restringe o espaço de busca sem cair na miopia de abordagens estritamente gulosas (*greedy*). Dessa forma, o *Beam Search* otimiza o processo de tomada de decisão no *AquaCrop*, avaliando o impacto de sequências futuras de ações para selecionar as ramificações de manejo que maximizam a soma das recompensas acumuladas.

Por outro lado, a função *Softmax* foi utilizada para converter os valores brutos calculados pelo modelo em uma distribuição de probabilidades sobre o conjunto de ações de irrigação disponíveis [Goodfellow et al. 2016]. Para incentivar a exploração do espaço de estados durante a fase de treinamento, aplicou-se um parâmetro de temperatura (τ) elevado na equação da *Softmax*. Essa calibração suavizou a distribuição de probabilidades, reduzindo a discrepância entre os valores das ações e permitindo que o agente experimentasse decisões menos prováveis — “aguçando sua curiosidade” e evitando a convergência prematura para subótimos locais. Essa abordagem probabilística serviu como uma linha de base (*baseline*) para comparar o impacto de uma escolha estocástica passo a passo em relação à busca estipulada e estruturada do *Beam Search*.

O restante deste artigo está organizado da seguinte maneira: a seção 2 apresenta a fundamentação teórica, abordando conceitos utilizados no trabalho, como os fundamentos do aprendizado por reforço e conceitos agrônômicos. A seção 3 discute e resume os trabalhos relacionados, contextualizando a proposta deste trabalho em relação a estudos

que aplicam *Reinforcement Learning* ao manejo hídrico. A seção 4 descreve o desenvolvimento do sistema, incluindo as etapas do experimento, como a proposta inicial com controle baseada em limiares, calibração de sensores, modelagem e treinamento do agente usando as políticas *Beam Search* e *Softmax*. Por fim, a seção 5 apresenta os resultados e considerações finais, discutindo as limitações do estudo e propostas de trabalhos futuros.

2. Fundamentação Teórica

2.1. Conceitos de Agricultura

O entendimento da dinâmica da água no solo é essencial para o manejo eficiente da água. Neste contexto, destacam-se conceitos fundamentais definidos pelo principal guia global da FAO para o cálculo de demanda hídrica das culturas, a FAO-56 [Allen et al. 1998].

Dentre esses conceitos estão a capacidade de campo (FC), que representa a capacidade de retenção de água em um solo após drenagem, e o ponto de murcha permanente (PWP), o qual é o nível mínimo de umidade a partir do qual as plantas não conseguem mais extrair água.

A diferença entre esses dois parâmetros define a água disponível no solo (*Total Available Water* – TAW). Este conceito é fundamental para encontrar o volume de água no solo, porque permite converter as medições relativas dos sensores em valores aproximados de volume, e com isso calcular o total de litros de água em um volume de solo.

Outro conceito relevante é a fração de depleção, que é usada no cálculo de água facilmente disponível para determinada cultura (*Readily Available Water* – RAW). Segundo [Allen et al. 1998], o RAW representa a parcela do TAW que uma planta pode consumir sem entrar em estresse hídrico e atrapalhar o desenvolvimento da cultura. Tal conceito é indispensável para encontrar a quantidade ideal de reposição de água a ser feita ao solo.

A evapotranspiração é o processo em que o solo e a planta perdem água para a atmosfera, e sua estimativa é um ponto fundamental para entender a demanda hídrica de determinada cultura. Segundo a FAO [Allen et al. 1998], o modelo de Penman-Monteith é o mais recomendado para esta finalidade. Ele combina variáveis climáticas, como umidade e temperatura do ar, radiação solar e velocidade do vento, para estimar a evapotranspiração de referência (ET₀). A evapotranspiração da cultura (ET_c) é obtida pela relação: $ET_c = ET_0 \times K_c$, onde K_c é o coeficiente da cultura, que varia conforme o tipo da planta e seu estágio de desenvolvimento.

A fórmula de Penman-Monteith é dada por:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

onde:

- ET_0 : evapotranspiração de referência (mm dia⁻¹)
- Δ : inclinação da curva de pressão de vapor (kPa °C⁻¹)
- R_n : radiação líquida na superfície (MJ m⁻² dia⁻¹)
- G : fluxo de calor no solo (MJ m⁻² dia⁻¹)
- γ : constante psicrométrica (kPa °C⁻¹)
- T : temperatura média do ar (°C)

- u_2 : velocidade do vento a 2 m de altura (m s^{-1})
- e_s : pressão de vapor de saturação (kPa)
- e_a : pressão real de vapor (kPa)
- $(e_s - e_a)$: déficit de pressão de vapor (kPa)

2.2. Aprendizado por reforço

O aprendizado por reforço (RL, do inglês *Reinforcement Learning*) é uma vertente do aprendizado de máquina na qual um agente é treinado para tomar decisões iterativas ao interagir com o ambiente. Diversamente da abordagem supervisionada, o agente não recebe respostas corretas explícitas, mas otimiza seu comportamento via tentativa e erro, dispensando conjuntos de dados rotulados [Pellegrini and Wainer 2007].

Neste cenário, o agente observa o estado atual do ambiente, executa uma ação e recebe uma recompensa de acordo com o impacto resultante dessa ação no ambiente. O objetivo principal é aprender a melhor estratégia de escolha de ações que maximize a soma das recompensas recebidas ao longo do tempo.

Este processo é modelado na forma de um Processo de Decisão de Markov (MDP, do inglês *Markov Decision Process*), conforme descrito por [Pellegrini and Wainer 2007], e representado pela tupla (S, A, T, R) , em que:

- S : é um conjunto de estados possíveis
- A : é um conjunto de ações que o agente pode executar
- $T : S \times A \times S \rightarrow [0, 1]$ é a função que dá a probabilidade de o sistema transitar do estado $s \in S$ para um estado $s' \in S$ após a execução de uma ação $a \in A$, ou seja, $T(s, a, s') = P(s'|s, a)$;
- R : é a função de recompensa que associa um valor de acordo com a execução da ação $a \in A$ no estado $s \in S$

No contexto deste trabalho, o espaço de estados S é representado pelo vetor de características $s = (U_{\text{solo}}, T_{\text{ar}}, U_{\text{ar}}, E_{\text{planta}})$, que consolida as condições instantâneas do ambiente. Para viabilizar a representação do Processo de Decisão de Markov e o mapeamento da política do agente, as variáveis contínuas foram discretizadas em intervalos finitos, de modo que cada combinação única desses parâmetros define um estado discreto e bem-definido do sistema.

As ações A não correspondem diretamente a volumes absolutos de irrigação, mas sim a fatores de multiplicação aplicados sobre um volume previamente definido pela fórmula de evapotranspiração da cultura (ET_c); assim, o conjunto de ações é definido como um conjunto de coeficientes. A função de transição $P(s'|s, a)$ é modelada por meio do simulador *AquaCrop*, que cria um ambiente de simulação agrícola após uma ação. E a função de recompensa $R(s, a)$ é definida por quão perto o nível de umidade do solo chegou ao nível ideal.

Um dos algoritmos mais utilizados para resolver problemas modelados como MDP é o Q-learning, um método de aprendizado por reforço que não depende do conhecimento prévio da dinâmica do ambiente. O Q-learning estima, para cada par estado-ação, um valor $Q(s, a)$ que representa a qualidade esperada daquela ação em um determinado estado, armazenando esses valores em uma tabela conhecida como *Q-table*. Ao longo do treinamento, o agente atualiza repetidamente sua *Q-table* a partir das interações com o ambiente, convergindo aos poucos para uma política que maximiza a recompensa acumulada [Pellegrini and Wainer 2007].

3. Trabalhos Relacionados

O presente trabalho fundamenta-se nas abordagens de Aprendizado por Reforço aplicadas ao manejo hídrico. Neste cenário se destacam trabalhos realizados por M. Chen et al. (2021) e Y. Chen et al. (2023, 2025) ao longo dos últimos anos.

No estudo de [Chen et al. 2021], foi proposto um modelo de redes Q-profundas (DQN) integrado a previsões meteorológicas para a irrigação de arroz. O método consistiu no treinamento do agente com base no balanço hídrico diário do solo, obtendo como resultado uma otimização significativa no aproveitamento da água da chuva e aumento na produtividade da cultura em relação aos manejos convencionais.

Posteriormente, [Chen et al. 2023] expandiram a abordagem ao propor um modelo de Aprendizado por Reforço Profundo focado em culturas de algodão. O método utilizou dados simulados para calibrar estratégias ótimas sob restrições hídricas severas. Como resultado, o sistema demonstrou alta capacidade de maximizar o rendimento produtivo reduzindo o consumo de água, alinhando-se a metas de sustentabilidade agrícola.

Recentemente, [Chen et al. 2025] avançaram na complexidade algorítmica ao aplicar a arquitetura *Distributional Actor-Critic* ao manejo de precisão. O método refinou a tomada de decisões ao mapear de forma probabilística os estressores ambientais e a umidade do solo. Os resultados evidenciaram um salto expressivo na eficiência do uso da água e na robustez do sistema para lidar com a escassez de recursos em regiões áridas.

Embora essa sequência de pesquisas valide a eficácia da inteligência artificial no ambiente agrícola e apresente um alto desempenho, tais abordagens demandam elevado poder computacional para executar arquiteturas pesadas (*Deep Q-Networks*, *Distributional Actor-Critic*), além de exigirem infraestrutura de *hardware* de custo elevado para a realidade do pequeno e médio produtor. O presente trabalho diferencia-se ao propor uma solução de baixa complexidade computacional — utilizando Aprendizado por Reforço aliado ao *Beam Search* para reduzir a carga de processamento — e baixo custo de implantação, viabilizado pelo uso do microcontrolador ESP32, sensores acessíveis e estruturas customizadas em impressão 3D.

No trabalho de [e Silva and Ítalo Rodrigo Monte Soares 2025], os autores desenvolveram um protótipo de irrigação automatizada utilizando o microcontrolador ESP32 integrado a sensores de umidade do solo para otimizar o uso da água no cultivo. O trabalho compartilha com o presente estudo o uso de infraestrutura IoT de baixo custo para a coleta de dados e a substituição do manejo manual por regras automatizadas.

4. Desenvolvimento

4.1. Experimento Inicial com Controle Baseado em Limiares

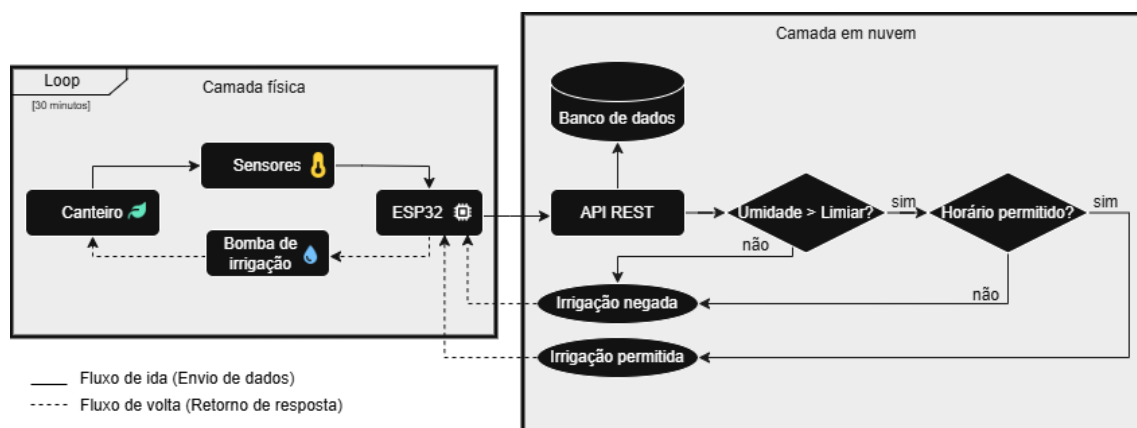


Figura 1. Diagrama do fluxo utilizado na automação da irrigação do canteiro

Durante o desenvolvimento inicial do sistema, a cultura alvo foi a alface americana, escolhida pela capacidade de produção contínua ao longo de praticamente todo o ano, e à sua elevada sensibilidade a falhas no manejo hídrico, essa característica impõe margens estreitas de tolerância hídrica, permitindo testar a capacidade do sistema em evitar episódios de seca ou encharcamento no solo. Foram utilizados sensores resistivos HL-69 pra medição da umidade do solo, um sensor DHT11 para aferição de temperatura e umidade do ar, e um microcontrolador ESP32. Essa seleção de hardware priorizou o baixo custo e a acessibilidade de mercado, viabilizando um protótipo funcional para a validação da arquitetura do sistema.

Para a proteção e suporte dos componentes no ambiente externo, foram desenvolvidas estruturas personalizadas em impressão 3D (Figuras 2 e 3).. Utilizou-se o filamento PETG (*Polietileno Tereftalato de Glicol*) por apresentar maior resistência térmica e mecânica em comparação ao PLA (*Ácido Polilático*), além de menor toxicidade em relação ao ABS (*Acrilonitrila Butadieno Estireno*) — que libera vapores nocivos como o estireno quando em altas temperaturas —, garantindo maior durabilidade ao protótipo a um custo acessível.



Figura 2. Abrigo meteorológico.



Figura 3. Proteção do módulo HL-69.

O sistema também utilizou uma API REST para processamento e um banco de dados relacional para armazenar os dados coletados, ambos hospedados em nuvem

O principal objetivo deste sistema foi a automatização de irrigação do canteiro para evitar o estresse hídrico, a economia de água não foi considerada, o sistema realizava verificações a cada 30 minutos, a tomada de decisões seguia regras de limiar de umidade e janelas de tempo fixas.

- **Critério de Umidade:** A irrigação somente era permitida se a média dos sensores de umidade atingiam certo valor pré estabelecido no banco de dados, que representava o limite mínimo de umidade. Esta limiar de umidade do solo foi definida através de um experimento de campo, que consistiu essencialmente em usar o sensor HL-69 pra mensurar o potencial hídrico do solo do canteiro em condição levemente úmida, o valor obtido foi estabelecido como o parâmetro de referência.
- **Janelas de tempo:** A irrigação também só era permitida se, além do critério de umidade, o horário estiver dentro de dois períodos diários, matutino (07:00 às 09:00) e vespertino (16:00 às 18:00), como a verificação era feita a cada 30 minutos, poderia ocorrer mais de uma irrigação dentro da janela de tempo permitida, se o critério de umidade também fosse satisfeito. Estes períodos foram escolhidos com o objetivo de maximizar a eficiência de absorção da planta de acordo com diretrizes publicadas pela Embrapa. Segundo [Marouelli et al. 2001], a irrigação no período matutino reduz a perda da água irrigada por evaporação. Já a restrição do período vespertino baseou-se nos alertas de [Lopes and Reis 2007], onde aponta que irrigações tardias provocam um acúmulo de água sob as folhas saias das plantas, e criam um microclima favorável pra proliferação de fungos, principalmente nas fases finais da planta.

Os testes de campo foram conduzidos em dois canteiros com as mesmas características de solo e adubação. Apenas um deles recebeu o sistema de monitoramento e irrigação automática (figura 4). O outro foi irrigado manualmente de acordo com a percepção visual e tátil da umidade do solo, sendo a quantidade inserida dinâmica (com volume máximo de até 5 litros por aplicação) e realizada estritamente nos turnos permitidos: matutino (7:00 às 09:00) e vespertino (16:00 às 18:00). (Figura 5).

Durante o período de cultivo (de 31 de outubro de 2025 até 31 de dezembro de 2025, onde a última planta foi colhida), pode-se observar que o canteiro equipado com o sistema inteligente apresentou um desenvolvimento maior das folhas das alfaces, e mesmo com uma planta sendo descartada, a pesagem total da produção foi maior.

Outro fenômeno relevante, foi que o canteiro convencional concentrou um maior número de plantas colhidas em períodos próximo à janela cronológica ideal, em contrapartida, o canteiro equipado com o sistema teve a maior parte de sua colheita realizada em uma data na qual a senescência e os sintomas de mela já estavam estabelecidos no campo. Este detalhe reforça o desempenho superior do canteiro automatizado, que, mesmo com condições desfavoráveis, superou a biomassa total do canteiro convencional.



Figura 4. Canteiro monitorado



Figura 5. Canteiro convencional

Tabela 1. Canteiro monitorado			
Linha	Planta	Data	Peso (g)
A	A1	27/12	362
A	A2	25/12	417
A	A3	27/12	279
A	A4	27/12	284
A	A5	27/12	*
A	A6	27/12	289
B	B1	27/12	417
B	B2	25/12	326
B	B3	19/12	458
B	B4	27/12	405
B	B5	21/12	474
B	B6	31/12	460
Soma			4171

*Planta não contabilizada devido à podridão.

Tabela 2. Canteiro convencional			
Linha	Planta	Data	Peso (g)
A	A1	31/12	208
A	A2	25/12	317
A	A3	19/12	468
A	A4	21/12	356
A	A5	21/12	277
A	A6	31/12	412
B	B1	27/12	357
B	B2	27/12	397
B	B3	27/12	276
B	B4	27/12	337
B	B5	21/12	368
B	B6	27/12	195
Soma			3968

Produtividade Total:

- **Canteiro Convencional:** 3.968g, 12 plantas colhidas, peso médio de 330,67 g/planta.
 - **Canteiro Monitorado:** 4.171 g, 11 plantas colhidas, peso médio de 379,18 g/planta.
- Ganho:** +203g (+5,1%) a favor do canteiro monitorado, mesmo com uma planta a menos.

Na fase final do ciclo, as alfaces de ambos os canteiros foram colhidas de forma escalonada e intercalada no período de 19/12/2025 a 31/12/2025. Essa estratégia de amostragem foi adotada por dois motivos principais: evitar o desperdício do material vegetal e permitir a avaliação temporal do desenvolvimento e da conservação das plantas no campo em momentos similares de desenvolvimento. Durante este tempo, foi observado que plantas colhidas após o dia 27/12/2025 entraram em fase de pendoamento, por excederem o período ideal de colheita, e sintomas de uma doença popularmente conhecida como "*Morredeira*" ou "*Mela*" nas folhas inferiores, conforme ilustrado na Figura 6.



Figura 6. Sintomas de "Morredeira" observados nas folhas inferiores de algumas alfaces

A ocorrência desta doença foi associada ao excesso de umidade no solo, causada pela estratégia inicial de irrigação do sistema. Embora esta estratégia tenha beneficiado as plantas mais jovens, ela se mostrou inadequada para plantas já desenvolvidas, que demandam menor frequência de irrigação.

O fator relevante foi a morfologia da cultura. As folhas largas da alface desenvolvida formam uma cobertura sobre o solo, e reduzem a incidência da luz solar e dificultam a evaporação da água abaixo da planta. Esse efeito contribui para a formação de um ambiente com umidade elevada próximo da raiz da planta, condição favorável para fungos e bactérias. Embora não tenha sido realizada a identificação do agente causador, os sintomas observados são compatíveis com condições de excesso hídrico.

Os sensores de umidade do solo não captaram diretamente essa umidade localizada, pois estavam posicionados entre as plantas, em locais que a evaporação ocorria com mais facilidade e não necessariamente próximos a raiz da planta. Assim, as leituras representavam a umidade média do canteiro, mas não refletiam o microambiente específico abaixo de cada planta, esta limitação contribuiu para que a irrigação ocorresse mesmo com a umidade elevada embaixo das plantas.

Estes resultados mostraram a necessidade de considerar os estágios de desenvolvimento da cultura alvo no manejo hídrico, porque regras fixas não são capazes de representar essa variabilidade. Para contornar este impasse, foi pensado em um modelo de inteligência artificial, que aprende o comportamento do solo a partir de dados históricos e se ajusta de acordo com a necessidade hídrica, considerando o estágio da planta e variáveis ambientais.

4.2. Aplicação de Inteligência artificial

O modelo de inteligência artificial proposto utiliza o algoritmo de Q-learning para corrigir eventuais imprecisões da previsão de evapotranspiração — decorrentes de variações não previstas nas condições meteorológicas — e tem o objetivo de prever a necessidade hídrica da cultura, mantendo o volume de água no solo dentro dos limites adequados. Para isso, evita-se que os níveis fiquem abaixo do limite da água facilmente disponível (RAW), adicionado por uma margem de 10% deste valor, resultando em um nível confortável para o desenvolvimento da planta, no qual pequenos erros não seriam tão críticos para a cultura. O sistema é modelado sob a estrutura de um Processo de Decisão de Markov (MDP), composto por:

- **Estado:** O cenário atual em que o canteiro se encontra, representado pela combinação do volume de água no solo, temperatura do ar, umidade do ar e estágio de desenvolvimento da cultura, categorizado em uma das quatro fases fenológicas: inicial (1), desenvolvimento (2), meio de estação (3) e fim de estação (4).
- **Ação:** A escolha do agente para ajustar o volume total de irrigação. A ação consiste em selecionar um entre cinco multiplicadores (-30% , -15% , 0% [manter volume], $+15\%$ e $+30\%$) que é aplicado diretamente sobre o volume total de água recomendado para a regra do canteiro.
- **Recompensa:** O *feedback* numérico fornecido pelo ambiente após a execução de uma ação. A função atribui recompensas maiores quanto mais próximo o volume ajustado estiver da necessidade hídrica ideal da cultura (volume alvo) e penalizações progressivas à medida que se afasta dela.
- **Função de transição:** A dinâmica de transição de estados é governada estocasticamente pelas simulações do software *AquaCrop*.

Para evidenciar as diferenças estruturais entre o modelo de decisão proposto e a abordagem inicial (Seção 4.1), a Tabela 3 sintetiza os principais parâmetros de cada sistema. O modelo baseado em limiares representa uma solução determinística e de baixa complexidade, que evidenciou as limitações discutidas anteriormente (regras fixas incapazes de considerar o estágio de desenvolvimento da cultura). Já o modelo de Q-learning incorpora variáveis adicionais ao processo decisório e ajusta sua resposta dinamicamente, a partir da política aprendida.

Tabela 3. Comparativo entre o modelo inicial (limiar fixa) e o modelo de RL (Q-learning)

Parâmetro	Modelo inicial (limiares)	Modelo RL (Q-learning)
Critério de decisão	Limiar fixo de umidade média	Estado multivariável ($U_{solo}, T_{ar}, U_{ar}, E_{planta}$)
Volume de irrigação	Fixo	Dinâmico *
Janela temporal	matutino/vespertino	N/A **
Estágio da cultura	Não considerado	Considerado (4 fases fenológicas)
Adaptabilidade	Nenhuma (regra estática)	Aprendida via interação com o ambiente

*O volume é calculado pela previsão de evapotranspiração (Penman-Monteith) e ajustado dinamicamente pela IA, por meio de um multiplicador entre -30% e $+30\%$.

**O horário de aplicação não é uma variável do modelo de RL, pois o cálculo da evapotranspiração (Penman-Monteith) já engloba as condições climáticas acumuladas em um período de 24 horas, independentemente do momento do dia em que a irrigação ocorre.

4.3. Calibração dos sensores capacitivos de umidade

A calibração dos sensores é uma etapa fundamental para garantir a precisão dos dados captados, e permitir a conversão dos valores brutos em volume de água no solo. Na segunda etapa deste trabalho, foram usados sensores capacitivos de umidade do solo para realizar experimentos, que funcionam retornando um valor de acordo com a variação da dielétrica do solo, em função do teor de água ao redor do sensor.

Os sensores foram calibrados considerando dois pontos de referência, o valor que foi captado com o sensor suspenso e sem contato com o solo, para não captar umidade (dry_reference_adc ou 0%) e o valor captado no solo no limite da FC (wet_reference_adc ou 100%), a terra usada foi coletada no mesmo canteiro em que os sensores vão ser inseridos.

Para determinar a capacidade de campo foi usado o método adaptado ao usado em [Gumieri et al. 2022], porém foram calibrados somente os extremos de umidade (0% e 100%) . Um pote cônico com 899 g do solo, foi saturado e deixado em drenagem livre até a estabilização do teor de água. Após 24 horas, foi medido novamente o peso do pote com terra, que totalizou 1.187 g, representando retenção de 288 g da água no solo, sendo o volume do pote 763cm^3 , a retenção foi de $37,7\%$ do volume total. O ponto de murcha permanente foi estimado com base em valores típicos da literatura para solos de franco-argilosos, onde o ponto de murcha corresponde, em média, a 50% a 60% da capacidade de campo. Por isso foi estipulado o valor aproximado de 20% do volume para o ponto de murcha permanente.

Esta informação mostra que em determinado volume deste solo, a retenção vai

ser de 37,7%, considerando um canteiro em forma de paralelepípedo, com dimensões de 0,15 m de altura, 2 m de comprimento e 0,6 m de largura, seu volume é dado por:

$$V = 0,15 \times 2 \times 0,6 = 0,18 m^3 \quad (2)$$

Convertendo para litros:

$$0,18 m^3 = 180 L \quad (3)$$

Com as informações coletadas, podemos afirmar que a FC deste canteiro feito com este solo é de aproximadamente 67,9 litros ($37,7\% \times 180 L$), tendo o PWP em 36 litros ($20\% \times 180 L$), seu TAW é de aproximadamente 31,9 litros ($FC - PWP$). Os valores de `wet.reference.adc`, ou 100%, foram coletados com o solo nesta condição.

Tendo os pontos de referência após calibração, os valores que são retornados por estes sensores no solo, representam uma porcentagem entre 0% e 100% de teor de água, valores que extrapolam 100% representam saturação.

4.4. Treinamento do Agente

O treinamento do agente de irrigação foi realizado em ambiente simulado, permitindo usar diferentes estratégias de irrigação e treinamento, sem necessidade de experimentação direta em campo.

Para viabilizar o uso do *AquaCrop* de modo a obter dados sequenciais, foi necessário adaptar sua lógica de simulação, originalmente orientada à execução completa do ciclo, sem suporte a iteração passo a passo. A estratégia adotada foi utilizar o arquivo de saída do *AquaCrop* que apresenta saídas diárias igualmente idênticas no fim da simulação, e ao executar a simulação com arquivo de irrigação preenchido, resulta em um arquivo com variáveis diárias, permitindo a extração da variável de interesse, no caso, a umidade do solo.

A reprodutibilidade dos estados simulados foi um fator fundamental para essa abordagem, porque um mesmo histórico de irrigação, em conjunto com as condições ambientais idênticas (temperatura e evapotranspiração de referência e características do solo) resulta em resultados diários idênticos. Desta forma, foi possível estruturar um mecanismo similar a uma simulação passo a passo.

Esta reprodutibilidade também possibilitou a aplicação do *Beam Search* no treinamento. A cada etapa são avaliadas múltiplas ações possíveis, representadas por cinco diferentes multiplicadores do volume de irrigação. Para cada ação, o sistema executa uma simulação no *AquaCrop*, e em seguida consulta o agente por meio de um *endpoint*, fornecendo o estado resultante desta ação. Em resposta, o modelo devolve o valor da recompensa associada, que é a relação entre o estado observado e o nível de umidade alvo, neste processo ocorre o aprendizado do modelo. Após as consultas de recompensas das ações, os estados com maiores recompensas acumuladas, limitadas ao tamanho do beam, são adicionados como históricos parciais, que serão expandidos nas próximas iterações. Desta forma, o algoritmo explora múltiplas trajetórias de irrigação de forma controlada, mantendo as promissoras ao longo da busca, enquanto treina o modelo a cada decisão avaliada. O uso do *Beam Search* no treinamento resulta em recompensas estáveis, porque seu treinamento é orientado justamente a isso, encontrar sempre os resultados onde a soma das recompensa vai ser maior.

A função de recompensa foi projetada de forma a fornecer sinais ao agente, implementando uma região de tolerância onde a recompensa é maximizada, além de uma sub-região com uma bonificação adicional. Esta estratégia é conhecida como *reward shaping*, e visa acelerar a convergência do agente, indicando regiões ideais nos estados. Por exemplo: considere um estado onde a umidade do solo apresenta um erro de 0.2 em relação ao valor alvo, como este valor está dentro da sub-região da tolerância, o agente recebe recompensa máxima, acrescida pela bonificação. Por outro lado, se o erro for maior, como 0.8, a recompensa passa a ser calculada pela função de penalização, sem acréscimos, porque está fora da região de tolerância.

4.4.1. Utilização de política Softmax

Inicialmente, o agente foi treinado utilizando uma política de seleção de ações baseada em *Softmax*, permitindo a exploração do espaço de ações. Ao longo do treinamento, a temperatura da exploração vai progressivamente decaindo, e o comportamento do agente vai se aproximando de uma política gulosa.

Em um treino isolado de um ciclo, o treino com política *Softmax* executa 75 simulações do *AquaCrop*, em contrapartida, o treino utilizando o algoritmo *Beam Search* executa 1831 no mesmo contexto, reforçando que, cada simulação gera uma linha de resultado correspondente a um dia avaliado. Esta abordagem depende da exploração das ações dentro dos estados para aprender sua política, então necessita de um grande volume de simulações.

Para analisar o comportamento do modelo que utiliza a política *Softmax* em um cenário com maior volume de treinamento, foram realizados 3 lotes de simulações com 30 ciclos cada. A escolha de 30 ciclos buscou equiparar o volume total ao treino do algoritmo *Beam Search*, que totalizou aproximadamente 24 ciclos. É importante ressaltar a diferença estrutural entre os treinos: enquanto no *Softmax* cada ciclo de treinamento corresponde a 75 dias (equivalente ao ciclo completo da cultura do alface), no *Beam Search* a dinâmica ocorre via árvore de busca com ramificações sucessivas baseadas na largura de beam de tamanho 3 ($k = 3$). Para ampliar a robustez e a adaptabilidade da política aprendida, a data de início de cada ciclo do *Softmax* foi avançada em um mês a cada iteração, expondo o agente a diferentes variações climáticas. Ao final de cada lote, realizou-se um ciclo de avaliação com início em janeiro de 2025 para mensurar o desempenho do modelo. O resultado do *Softmax* indicou aprendizado progressivo via aumento de recompensas, porém mantendo-se inferior ao *Beam Search*. Os testes demonstraram convergência mais rápida e estável no *Beam Search*, que atingiu uma recompensa média de 8,60 após um ciclo, contra 6,55 obtidos pelo *Softmax* ao fim de 90 ciclos. Contudo, por ser um algoritmo de busca heurística local com largura definida em 3, o *Beam Search* consolida resultados 'confortáveis' e seguros precocemente ao podar trajetórias promissoras que dependem de ações subótimas no curto prazo. Por outro lado, o *Softmax* mantém a exploração ativa ao longo dos 75 dias de cada ciclo, o que, embora exija um volume substancialmente maior de treino para convergir, preserva o potencial de encontrar o ótimo global a longo prazo.

Como mostra a figura 7, a cada lote de treinamento aperfeiçoa a política *Softmax*, que apresenta média de recompensas crescentes

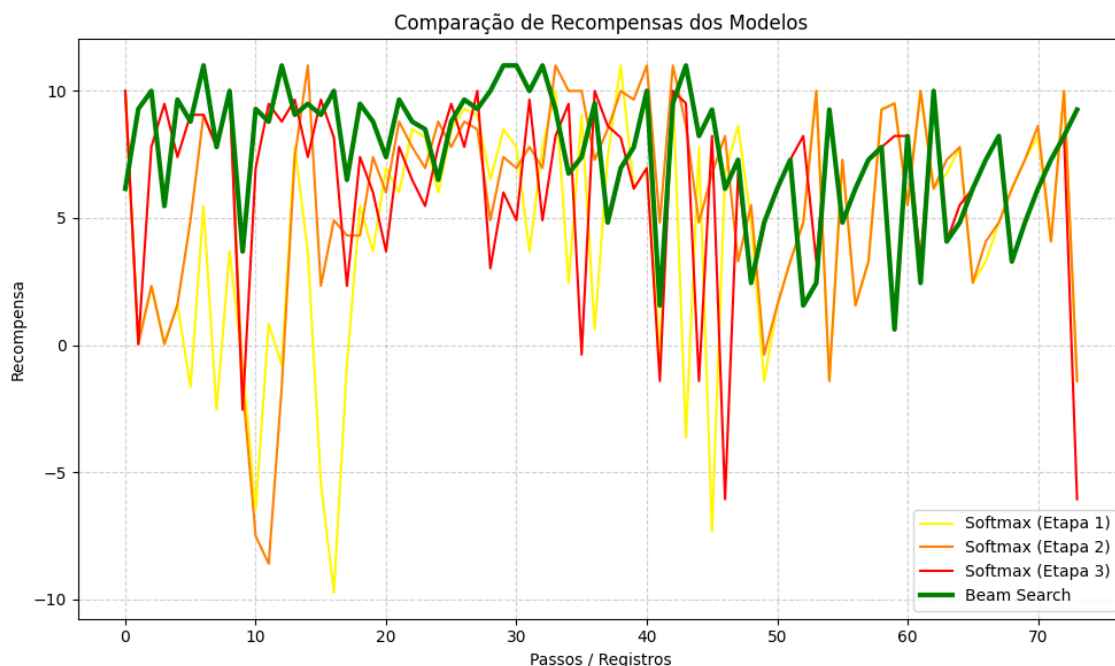


Figura 7. Comparação das recompensas atingidas pelos modelos, durante as simulações de avaliação

5. Considerações Finais

Este trabalho apresentou o desenvolvimento e a avaliação de um sistema de irrigação inteligente, focado no manejo hídrico otimizado na cultura da alface. Os experimentos iniciais em campo validaram a eficiência do monitoramento contínuo e da coleta de dados em tempo real no canteiro, mas também mostraram que sistemas baseados em regras estáticas levariam a perda de recursos valiosos, como água e no comprometimento da integridade da produção, devido a fungos.

Contudo, o algoritmo *Beam Search* demonstrou nítida superioridade sobre a política *Softmax*: por adotar uma busca heurística conservadora com largura de *beam* 3, o agente evitou punições severas desde as primeiras iterações e alcançou uma recompensa média superior (8,60) com um volume de treinamento substancialmente menor.

Entretanto, é importante destacar que os resultados apresentados foram obtidos exclusivamente em ambiente simulado, utilizando o modelo *AquaCrop* como função de transição do sistema. Dessa forma, fatores presentes em cenários reais, como variabilidade espacial do solo, falhas de sensores e ruídos ambientais, não foram considerados neste estudo.

Como trabalho futuro, propõe-se a validação dos modelos em um canteiro real com a integração de microcontroladores e sensores para a coleta de dados do solo, como feito neste trabalho. Este experimento permitirá avaliar o impacto da política aprendida em condições reais, sua capacidade de adaptação e a discrepância entre sua aplicação nas simulações e na realidade.

Referências

- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). Crop evapotranspiration: Guidelines for computing crop water requirements. Technical Report 56, Food and Agriculture Organization of the United Nations (FAO), Rome.
- Chen, M., Cui, Y., Wang, X., Xie, H., Liu, F., Luo, T., Zheng, S., and Luo, Y. (2021). A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. *Agricultural Water Management*, 250:106838.
- Chen, Y., Lin, M., Yu, Z., Shen, J., Xu, D., and He, L. (2025). Enhancing cotton irrigation with distributional actor-critic reinforcement learning. *Agricultural Water Management*, 307:109152.
- Chen, Y., Yu, Z., Han, Z., Sun, W., and He, L. (2023). A decision-making system for cotton irrigation based on reinforcement learning strategy. *Agronomy*, 14(1):11.
- Da Silva, C. L., Bassi, N. S. S., and Rocha Jr, W. F. D. (2016). Technologies for rational water use in Brazilian agriculture. *Ambiente e Agua - An Interdisciplinary Journal of Applied Science*, 11(2):239.
- e Silva, S. V. S. and Ítalo Rodrigo Monte Soares (2025). Irrigação automatizada por esp 32. *Revista Ibero-Americana de Humanidades, Ciências e Educação - REASE*, 11(11):3711–3723.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press, Cambridge, MA, USA.
- Gumieri, P. H. et al. (2022). Calibração de sensores capacitivos para a estimativa da umidade do solo. *Revista Cultivando o Saber*, 15(2):116–126.
- Liu, J., Chang, F., Yang, J., Jie, X., Lu, C., Wang, C., Xie, L., Ma, L., and Su, H. (2025). Deep Reinforcement Learning for irrigation optimization: Advantages, opportunities, and challenges. *Agricultural Water Management*, 322:110030.
- Lopes, C. A. and Reis, A. (2007). Doenças da alface. Circular Técnica 49, Embrapa Hortaliças, Brasília, DF.
- Lowerre, B. T. (1976). *The HARPY speech recognition system*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, USA.
- Marouelli, W. A., Silva, W. L. C., and Silva, H. R. (2001). Manejo da irrigação em hortaliças. Manual técnico, Embrapa Hortaliças, Brasília, DF.
- Pellegrini, J. and Wainer, J. (2007). Processos de decisão de markov: um tutorial. *RITA (Revista de Informática Teórica e Aplicada)*, 14.
- Russell, S. and Norvig, P. (2020). *Artificial Intelligence: A Modern Approach*. Pearson, Hoboken, NJ, USA, 4th edition.
- Salman, M., García Vila, M., Fereres, E., Raes, D., Steduto, P., Heng, L., De Iannoy, G., Wellens, J., Foster, T., Busschaert, L., Bechtold, M., and de Roos, S. (2025). AquaCrop on the ground – Model applications for sustainable agricultural water management. Technical Report 48, FAO, Rome.