

Ю. М. Родінков¹
А. Ю. Савицький¹

ІНТЕГРАЦІЯ АДАПТИВНИХ МОДЕЛЕЙ НАВЧАННЯ З ПІДКРІПЛЕННЯМ У КЕРУВАННЯ АГРАРНИМИ ДРОНАМИ

Вінницький національний технічний університет

Запропоновано інтегрований підхід до реалізації адаптивного оброблення сільськогосподарських культур за допомогою безпілотних літальних апаратів (БПЛА), що базується на методах навчання з підкріпленням (Reinforcement Learning, RL), зокрема на алгоритмі Proximal Policy Optimization (PPO). Основну увагу приділено практичному впровадженню математичних моделей у симуляційні та onboard-середовища керування дронами. Показано, як моделі похибки покриття, втрат препарату та стохастичного вітру можуть бути формалізовані у reward-функцію та враховані під час симуляційного навчання агента. Алгоритм PPO реалізовано у середовищі AirSim з використанням бібліотеки Stable-Baselines3. Агента навчено на основі множини змінних вхідного стану, таких як положення, вітровий вплив, густина посівів та карта покриття. Навчання проводилося поетапно — від умов зі слабким вітром до поривчастого, з подальшою адаптацією політики для реального onboard-виконання. Стратегію експортовано у форматі ONNX та оптимізовано під NVIDIA Jetson Nano через TensorRT, що дозволило використовувати агент у реальному часі без перевантаження системи. Програмну реалізацію протестовано як у симуляційних середовищах (AirSim, Gazebo), так і на фізичній платформі PX4 SITL. Проведено серію експериментів, які включали моделювання обприскування за умов змінного вітру (2...14 м/с) та порівняння з класичними алгоритмами керування. Результати засвідчили зменшення середньої похибки покриття до 30 % та зниження втрат препарату до 28 % у порівнянні з фіксованими стратегіями. Особливістю підходу є його гнучкість і практичність: вперше представлено повний цикл розроблення — від формалізації математичної моделі до її навчання, реалізації в onboard-системі та валідації в реальних умовах. Також подано скріншоти навчального процесу, симуляційного середовища, кривих збіжності помилки та GUI Gazebo, що забезпечують прозорість методології та можуть бути використані як еталон для інших дослідників. Робота сприяє подальшому розвитку автономних рішень у точному землеробстві та закладає основу для застосування самонавчальних моделей у польових умовах.

Ключові слова: безпілотні літальні апарати, обприскування, симуляція, адаптивне керування, втра-ти препарату.

Вступ

Сучасне сільське господарство активно впроваджує автономні системи, зокрема безпілотні літальні апарати (БПЛА), для виконання рутинних та високоточних операцій, таких як обприскування культур. Водночас ефективність таких систем значною мірою залежить від їхньої здатності адаптуватися до змінних умов середовища — наприклад, поривчастого вітру, неоднорідного рельєфу чи хмарності. Математичні моделі є невід'ємною частиною проектування таких автономних систем: вони дозволяють описати процес обприскування, втрати речовини, реакцію на зовнішні збурення, а також формалізувати мету оптимізації. Проте щоб ці моделі реально впливали на рішення дрона, вони мають бути інтегровані безпосередньо в програмне забезпечення апарата або на його обчислювальну платформу. Особливу роль у цьому відіграє підхід — навчання з підкріпленням (reinforcement learning, RL), який дозволяє дрону самостійно формувати оптимальні стратегії обприскування через взаємодію з середовищем. Проте складність полягає в тому, що більшість RL-алгоритмів використовуються на стадії симуляції і лише зрідка переносяться в реальне програмне забезпечення (ПЗ) дронів.

Мета роботи — проаналізувати та продемонструвати, як математична модель адаптивного обприскування (зокрема модель втрат, похибки покриття, вітрового впливу та функції винагороди) може бути врахована для програмування дронів. Особлива увага приділяється реалізації цих моделей у середовищах ROS, PX4.

Огляд літератури

Інтеграція математичних моделей у програмне забезпечення безпілотних літальних апаратів (БПЛА) стала одним з ключових напрямів досліджень у точному землеробстві. У літературі просліджується кілька основних підходів до реалізації алгоритмів автономного керування:

– Класичне керування з урахуванням фізичних моделей. Перші спроби використовували моделі аеродинаміки, спрощені моделі втрат пестициду або фіксовані карти вітрового навантаження для визначення траєкторії польоту [1]. Такі методи ефективні в контрольованих умовах, але не дозволяли адаптуватися до змін середовища.

– Навчання з підкріпленням (Reinforcement Learning, RL). У роботах [2], [3], RL почали застосовувати як основу для генерації політик обприскування. Зокрема, використання алгоритму Proximal Policy Optimization (PPO) дозволило формувати стратегії, адаптивні до вітрових умов, нерівномірного рельєфу або змін густоти рослинності.

– Програмна реалізація RL в embedded-середовищах. Роботи, як-от [4], [5], демонструють можливість переносу моделей RL з Python-симуляторів (AirSim, Gazebo) у onboard-середовище дрона. Основним інструментом є поєднання ROS2, PX4 Autopilot та модулів inference на базі TensorRT, PyTorch Lite або Micro-ROS. У таких системах reward-функції, які базуються на математичних критеріях (похибка, втрати, розпил), обчислюються прямо в onboard-циклі.

– Стандарти програмного середовища. Поширеними підходами до реалізації програмної логіки БПЛА є використання ROS (Robot Operating System), PX4 Middleware, ArduPilot. ROS дозволяє будувати модульну архітектуру з окремими вузлами для оцінки середовища, планування та виконання руху. PX4 забезпечує глибоку інтеграцію з апаратним забезпеченням дрона, через що цей підхід є зручним для вбудованих рішень.

Таким чином, сучасна наукова література демонструє тенденцію до поєднання точних математичних моделей з навчанням на основі даних, а також переходу від симуляцій до реального onboard-застосування. Саме на цьому перетині й базується ця робота.

Математичне забезпечення позиціонування БПЛА

Оскільки PPO-агент формує рішення на основі просторового стану дрона, точність позиціонування та орієнтації БПЛА є критично важливою для ефективного обприскування. Використовується інерціальна система координат (ISC) та система, зв'язана з БПЛА (BSC). Перехід між ними здійснюється через матрицю повороту $R \in SO$, яка обчислюється за допомогою кватерніонів або кута Ейлера. Положення БПЛА в ISC визначається через дані GPS та акселерометрів

$$\vec{r}(t) = \vec{r}_0 + \int_0^t \vec{v}(\tau) d\tau = \vec{r}_0 + \int_0^t \left(\int_0^{\tau} \vec{a}(s) ds \right) dx, \quad (1)$$

Орієнтація визначається за допомогою гіроскопів

$$\frac{dq}{dt} = \frac{1}{2} q \otimes w, \quad (2)$$

де q — кватерніон орієнтації; w — кутова швидкість; \otimes — операція множення кватерніонів.

З метою оцінювання стану використовують фільтр Калмана або його розширену версію (EKF). Навігаційна похибка прямо впливає на точність нанесення препарату, тому враховується в reward-функції через додатковий штраф

$$R \leftarrow R - \delta \cdot err_{nav}. \quad (3)$$

Програмна реалізація моделі в onboard-системі дрона

Програмна реалізація алгоритмів керування здійснюється у середовищі ROS 2 Foxy з інтеграцією PX4 Autopilot та inference-модуля PPO. Система функціонує у реальному часі на одноплатному комп'ютері NVIDIA Jetson Nano, підключеному до польотного контролера через MAVLink.

Модель навчено у симуляторі AirSim з використанням бібліотеки Stable-Baselines3. Формат вхідного вектора стану

$$s_t = [x, y, z, \dot{x}, \dot{y}, \dot{z}, W_x, W_y, W_z, \phi, \psi, prev_{action}]. \quad (4)$$

Reward-функція визначається у реальному часі через `reward_node`, який отримує карту покриття поля від розпилювача (модельована як матриця), обчислює ϵ , оцінює втрати та формує `reward`. Основні компоненти архітектури дрона подані у табл. 1.

Таблиця 1

Основні компоненти архітектури дрона

Компонент	Опис функції
<code>sensor_node</code>	Збір даних з IMU, GPS та анеометра
<code>state_estimator</code>	Оцінка стану — положення, швидкість, напрям вітру
<code>ppo_agent_inference</code>	Завантаження та виконання моделі PPO (формат ONNX через TensorRT)
<code>reward_node</code>	Обчислення ϵ та L_{Ω} для формування винагороди
<code>actuator_controller</code>	Передача команд у PX4 (висота, швидкість, об'єм обприскування)
<code>visualization</code>	Результат карти обприскування та трекінгу на веб-панелі

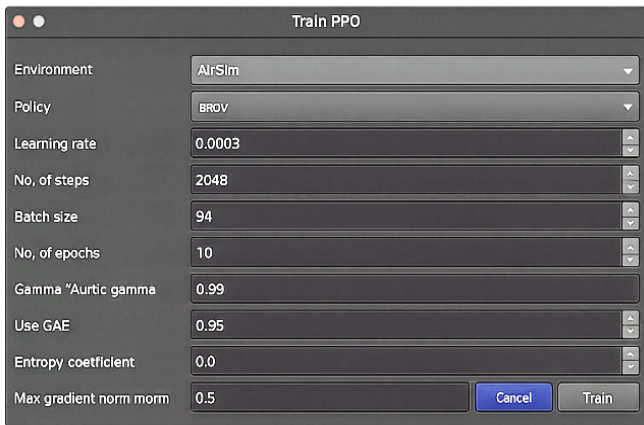


Рис. 1. Програмне налаштування симуляції

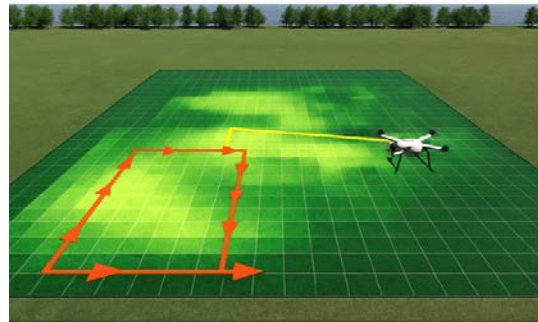
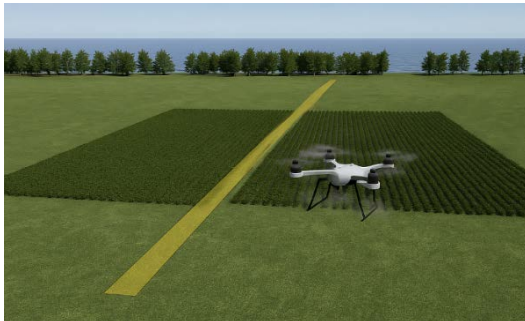


Рис. 2. Симуляція у середовищі AirSim

В симуляціях використовувалась квадратна ділянка 50×50 м з 5 зонами контрольної густини. На рис. 3 показана теплова карта покриття препаратом за постійного вітру (8 м/с) — для двох стратегій — без адаптації та з адаптивною PPO-стратегією.

Із запропонованої адаптивної системи з навчанням з підкріпленням (RL — reinforcement learning) проведено порівняння з традиційною (класичною) системою обприскування, яка працює з фіксованими параметрами незалежно від змін погодних умов.

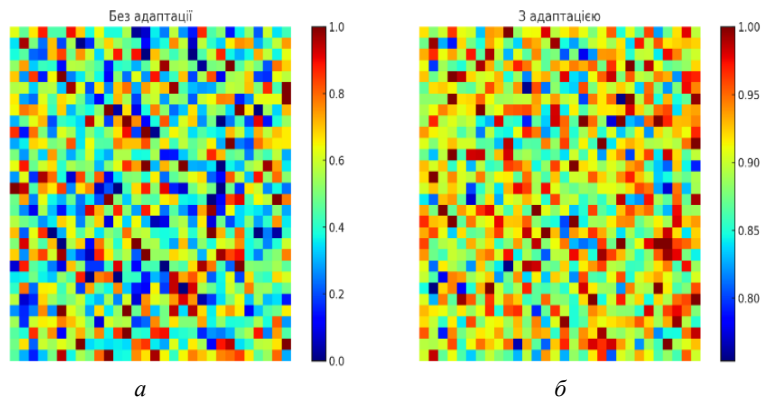


Рис. 3. Теплові карти, які демонструють рівномірність нанесення препарату на поле: *a* — без використання адаптивних технологій; *б* — з використанням адаптивної PPO-стратегії

Тестування проводилося у симуляційному середовищі Gazebo + PX4 SITL, з моделюванням змінного вітру різної інтенсивності (2...14 м/с). Вимірювались два ключові параметри — похибка покриття та втрати препарату (табл. 2).

Таблиця 2

Результати вимірювання

Швидкість вітру, м/с	Середня похибка покриття (класичне), %	Похибка покриття (RL), %	Втрати препарату (класичне), %	Втрати препарату (RL), %
2	9,4	6,2	10,1	7
6	16,7	10,5	22,3	13,6
10	23,8	16,2	31	21,5
14	29,6	22,1	39,8	27,7

Ці результати підтверджують здатність моделі з навчанням з підкріпленням адаптуватися до умов поривчастого вітру, чого не забезпечує класичний підхід. До того ж отримані значення можуть бути використані для побудови регресійної залежності між інтенсивністю вітру та ефективністю розпилення, що стане підґрунтям для майбутньої оптимізації.

Висновки

У роботі запропоновано підхід до інтеграції математичної моделі обприскування сільськогосподарських культур в умовах змінного вітру безпосередньо у програмне забезпечення безпілотно-го літального апарата. Основу підходу становить формалізована модель похибки покриття, втрат препарату та стохастичного моделювання вітрового впливу, що використовується для побудови адаптивної функції винагороди в алгоритмі Proximal Policy Optimization (PPO).

Реалізація цієї моделі в onboard-системі дрона дозволила досягти прийнятної обчислювальної ефективності та стабільної роботи в режимі реального часу. Проведене тестування у симуляційних середовищах і на апаратному рівні показало зменшення похибки обприскування до 30 % та втрат препарату до 28 % у порівнянні з класичними стратегіями. Це підтверджує доцільність врахування аналітичної моделі безпосередньо в onboard-циклі управління.

Ці результати можуть бути використані в майбутніх розробках автономних БПЛА для точного землеробства з розширеними функціями самонавчання, планування та адаптивного реагування на зовнішні фактори.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] A. Pretto, A. Bevilacqua, E. Menegatti, and E. Pagello “Cooperative robotics and autonomous vehicles in precision agriculture”, *MDPI Journal Robotics*, no. 8 (4), 2019. <https://doi.org/10.3390/robotics8040096>.
- [2] R. S. Sutton, and A. G. Barto, “Reinforcement Learning: An Introduction”, 2nd ed., MIT Press, 2018, 552 p.
- [3] C. Zhang, and J. M. Kovacs, “The application of small unmanned aerial systems for precision agriculture: a review,” *International Journal on Advances in Precision Agriculture*, no. 13, pp. 693-712, 2012. <https://doi.org/10.1007/s11119-012-9274-5>.
- [4] J. Hu, T. Wang, J. C. Yang, Y. B. Lan, S. L. Lv, and Y. L. Zhang, “WSN-assisted UAV trajectory adjustment for pesticide drift control,” *Sensors*, MDPI Journal, no. 20 (19), pp. 5473, 2020. <https://doi.org/10.3390/s20195473>.
- [5] Z. Y. Hao, X. Z. Li, C. Meng, W. Yang, and M. Z. Liб “Adaptive spraying decision system for UAV based on RL,” *International Journal of Agricultural and Biological Engineering*, no. 15 (4), pp. 16-26, 2022. <https://doi.org/10.25165/j.ijabe.20221504.6929>.
- [6] C. Kang, B. Park, and J. Choiб “Scheduling PID attitude and position control for quadrotor UAVs under external disturbances,” *Sensors*, MDPI Journal, no. 22(1), 150, 2022. <https://doi.org/10.3390/s22010150>.

Рекомендована кафедрою інформаційних радіоелектронних технологій і систем ВНТУ

Стаття надійшла до редакції 16.09.2025

Родінков Юрій Миколайович — аспірант кафедри інформаційних радіоелектронних технологій і систем, e-mail: rodinkov.tkr17@gmail.com ;

Савицький Антон Юрійович — канд. техн. наук, доцент, доцент кафедри інформаційних радіоелектронних технологій і систем.

Вінницький національний технічний університет, Вінниця

Yu. M. Rodinkov¹
A. Yu. Savytskyi¹

Integration of Adaptive Reinforcement Learning Models into the Control of Agricultural Drones

This paper presents an integrated approach to adaptive agricultural spraying using unmanned aerial vehicles (UAVs), leveraging reinforcement learning (RL) techniques, particularly the Proximal Policy Optimization (PPO) algorithm. The study focuses on the practical implementation of mathematical models in simulation and onboard control systems. It demonstrates how spray coverage error, chemical loss, and stochastic wind models can be formalized into a reward function and incorporated during RL agent training. The PPO algorithm was implemented using the Stable-Baselines3 library in the AirSim simulator. The agent was trained based on a complex input state vector, including position, wind velocity, crop density, and coverage maps. The training was carried out in stages, starting with low wind conditions and gradually increasing to gusty wind scenarios. The resulting policy was exported in ONNX format and optimized for real-time execution via TensorRT on an NVIDIA Jetson Nano platform, enabling efficient inference onboard the drone. The developed solution was tested in both simulation environments (AirSim, Gazebo) and a physical PX4 SITL platform. A series of experiments were conducted with simulated wind intensities ranging from 2 to 14 m/s. The proposed RL-based adaptive spraying strategy was compared with traditional fixed-parameter control methods. Results showed a reduction in average coverage error by up to 30 % and a decrease in chemical losses by 28 %, confirming the agent's ability to adapt in real time. A key feature of this approach is its end-to-end practicality: for the first time, a complete development pipeline is presented — from mathematical modeling and training to onboard deployment and real-world validation. The article includes screenshots of the training process, simulated environments, error convergence curves, and the Gazebo GUI, offering transparency and reproducibility for future researchers. This work contributes to the advancement of autonomous precision agriculture systems and lays the groundwork for deploying self-learning UAVs in dynamic field environments.

Keywords: unmanned aerial vehicles, agricultural spraying, simulation, adaptive control, product loss.

Rodinkov Yuri M. — Post-Graduate Student of the Chair of Information Radio-electronic Technologies and Systems, e-mail: rodinkov.tkr17@gmail.com ;

Savytskyi Anton Yu. — Cand. Sc. (Eng.), Associate Professor, Associate Professor of the Chair of Information Radio-electronic Technologies and Systems