

Міркевич О.М.

Національний університет харчових технологій

РОЗРОБКА МОДЕЛІ ДЛЯ ПРОГНОЗУВАННЯ ВОЛОГОСТІ ЖОМУ В ЖОМОСУШИЛЬНОМУ КОМПЛЕКСІ ЦУКРОВОГО ВИРОБНИЦТВА

У статті розглянуто побудову та виробничу валідацію моделі машинного навчання для прогнозування вологості жому в жомосушильному комплексі цукрового виробництва з горизонтом +30 хв. Актуальність зумовлена потребою в проактивному керуванні якістю та енергоефективністю сушіння, де традиційні емпіричні налаштування і процедурні правила часто не враховують інерційність процесу та перехідні режими, що знижує стабільність показників. Запропонований підхід спирається на виробничу телеметрію з дискретністю 30 хв і причинно коректні часові ознаки, зокрема лаги 1–6 кроків, короткі ковзні середні $MA(2,4,6)$ та градієнти $(2,4,6)$, які відбивають динаміку матеріального навантаження і теплового контуру. Базову модель *Random Forest* оцінено за схемою *TimeSeriesSplit* із двома фолдами з подальшим агрегуванням прогнозів на всій тестовій вибірці. Отримано показники точності $RMSE$ 0,229%, MAE 0,180%, R^2 0,025, а також $nMAE \approx 1,40\%$ і $nRMSE \approx 1,78\%$. Покриття малих похибок становить $|e| \leq 0,2\%$ у $\approx 56,5\%$ випадків і $|e| \leq 0,5\%$ у $\approx 98,9\%$ спостережень. Порівняння з найвими еталонами показує, що на квазістаціонарних ділянках персистентний прогноз Y_t і $MA(3)$ мають нижчі $RMSE$ та MAE , тоді як у перехідних режимах перевагу отримує *ML*-модель, що обґрунтовує гібридну політику застосування. Профіль важливостей підтверджує домінування груп ознак, пов'язаних із сухими речовинами на вході та температурою на виході, а короткі похідні $MA2$, $lag1$ і $grad2$ надають моделі чутливості до швидких змін. Запропоноване рішення формує проактивні рекомендації оператору з переоцінкою кожні 30 хв, зменшує перерегулювання і підвищує *time-in-band* у реалістично визначеному коридорі якості, створюючи підґрунтя для інтеграції у *SCADA* або *PLC* інтерфейс, розгортання швидкого режиму з горизонтом +15 хв і впровадження моніторингу дрейфу з плановим перенавчанням моделі.

Ключові слова: модель, жомосушильний комплекс, часові ознаки, прогнозування вологості, машинне навчання, *Random Forest*.

Постановка проблеми. Цукрова промисловість є одним із провідних сегментів харчової промисловості України. Характерним побічним продуктом її виробництва виступає буряковий жом. Ефективність операцій його сушіння визначально залежить від вологості сировини на вході до сушарки. Підвищення продуктивності та енергоефективності керування можливе за рахунок регулювання подачі палива в топку залежно від поточної вхідної вологості, однак побудова такого контуру ускладнюється браком надійних промислових засобів високоточного безперервного (*in-line*) вимірювання вологості потоку.

За відсутності прямих вимірювань на практиці орієнтуються на непрямий індикатор інтенсивності сушіння, тобто температуру димових газів на виході з барабана, яка опосередковано відображає різницю вологості між входом

і виходом. Процес має виражену інерційність, тому вплив змін вхідних параметрів позначається на якості продукту із запізненням приблизно на 30–40 хвилин. Поєднання непрямого зворотного зв'язку та динамічної запізненості зумовлює реактивний характер керування, коли корекції вносяться пізніше за збурення вологості сировини, зміни тягодуттєвих режимів і теплотворної здатності палива. У результаті зростає ризик перерегулювання і коливань вологості, збільшуються питомі енерговитрати через роботу з додатковими запасами, посилюється залежність якості від індивідуальних дій оператора, а діагностика причин відхилень ускладнюється.

З огляду на наведені обмеження, дослідження зосереджене на розробленні та інтеграції моделі машинного навчання, яка, спираючись на технологічну телеметрію та причинно коректні часові ознаки, прогнозує вихідну вологість

на горизонті близько 30 хв і перетворює цей прогноз на стандартизовані настанови щодо керування насамперед подачею жому. За потреби модель узгоджує це з мінімальними корекціями теплотехнічних і тягодуттєвих режимів з урахуванням технологічних обмежень (час перебування, рівномірність завантаження, допустимі навантаження приводу). Запропонований підхід переводить систему з реактивного у передбачувальний режим, зменшує амплітуду коливань вологості, підвищує частку часу в заданому коридорі якості, знижує питомі енерговитрати та скорочує залежність результатів від індивідуальних рішень оператора.

Аналіз останніх досліджень і публікацій. Сушіння бурякового жому належить до найенергоємніших операцій бурякоцукрового виробництва. На цій стадії формується значна частка питомих витрат палива та електроенергії, а коливання вологості на виході безпосередньо впливають на гранулювання, умови зберігання та логістичні процеси [16; 25]. У промисловій практиці переважають барабанні сушарки прямого типу з безпосереднім підігріванням димовими газами, і якість керування такими установками визначається паливно-повітряним режимом, параметрами тягодуттєвої групи, вологістю жому на вході та часом перебування матеріалу в барабані [9; 18; 24].

Динаміка роторної сушарки має виражену інерційність, тому зміни вмісту сухих речовин у жомі та інтенсивності його подачі, а також параметрів теплогенератора і тяги відбиваються на вологості жому на виході зі значним запізненням. Результати моделювання і чисельних симуляцій підтверджують, що навіть спрощені описи повинні враховувати розподіл часу перебування, перебіг теплообміну і масообміну вздовж барабана та внутрішні збурення, інакше керування набуває реактивного характеру [6; 7; 12]. Ранні прикладні дослідження у харчовій та цукровій галузях свідчать, що за обмеженого набору вимірюваних параметрів можна досягати прийнятної стабільності процесу, однак це потребує енергетичних резервів і посилює залежність від досвіду персоналу [5; 6].

Зворотний зв'язок традиційно формували на основі непрямих індикаторів інтенсивності сушіння, насамперед температури димових газів на виході барабана, у поєднанні з періодичним лабораторним контролем вологості. Такий сигнал забезпечує оперативність, проте не є прямою мірою майбутньої якості у момент ухвалення

рішень, тому корекції запізняються і зростає ймовірність перерегулювання [9; 16]. Сучасні мікрохвильові вологоміри для сипких та харчових матеріалів демонструють придатність інлайн вимірювання з достатньою точністю і робастністю, що відкриває можливість замикати контури на сигнал, безпосередніше пов'язаний із цільовими показниками якості, і водночас зменшувати потребу в енергетичних резервах [14; 23].

Розвиток алгоритмів поступово перейшов від цифрових регуляторів і евристичних правил до підходів прогнозного керування на основі моделей. Урахування в моделях транспортних запізнень, обмежень і міжканальних взаємодій зменшує коливання та наближає роботу до технологічних меж безпеки [5; 7; 22]. Паралельно інтенсивно розвиваються методи, що ґрунтуються на даних. Огляди та прикладні дослідження демонструють застосування ML і нейронних мереж для короткострокового прогнозування вологості та проактивних корекцій до того, як відхилення стає спостережним на виході [15; 17; 29]. У виробничих рішеннях дедалі частіше поєднують обидві парадигми: прогноз формують інструменти ML або глибинного навчання, а дотримання обмежень і баланс компромісів забезпечує модельно передбачувальне керування MPC [29; 15].

З позицій енергоефективності під час сушіння димовими газами наукові джерела послідовно підкреслюють ключову роль утилізації тепла та рециркуляції, оптимізації коефіцієнта надлишку повітря і ретельного налаштування тягодуттєвого обладнання. Такі заходи надійно знижують питомі енерговитрати, і цей ефект зберігається незалежно від обраної стратегії керування [9; 13; 16; 18]. Поряд із цим виробнича практика свідчить, що в уже експлуатованих установках найреалістичніший резерв економії пов'язаний із переходом від реактивного до передбачувального керування якістю, оскільки зменшення розкиду вологості та робота без додаткових запасів безпосередньо скорочують витрати палива [15; 17].

Зростає інтерес до цифрових двійників як до безпечного випробувального середовища для аналізу гіпотетичних сценаріїв, прогнозування моменту завершення сушіння відповідно до цільових показників якості та попереднього налагодження політик керування перед промисловим упровадженням [20]. У контексті сушіння бурякового жому поєднання цифрового двійника з короткостроковим прогнозом вологості на основі методів машинного навчання слугує

основною підсистеми підтримки прийняття рішень і природно інтегрується в наявну PLC/SCADA інфраструктуру.

Мета дослідження полягає у розробленні, обґрунтуванні та експериментальній перевірці моделі машинного навчання для прогнозування вологості жому в жомосушильному комплексі цукрового виробництва. Модель формує прогноз вихідної вологості на горизонті +30 хв на основі технологічної телеметрії та причинно коректних часових ознак і перетворює цей прогноз на проактивні операторські рекомендації з пріоритетом керування подачею жому. Інтегрована до підсистеми підтримки прийняття рішень реалізація моделі усуває типові вади традиційних схем керування, зокрема реактивність, запізнення та потребу в ручному підлаштуванні, і забезпечує стабілізацію якості за нижчих енерговитрат. Результати ефективності підтверджені експериментально на виробничих даних.

Дані та попередня обробка. У дослідженні використано телеметричні дані жомосушильного комплексу з рівномірним інтервалом спостережень 30 хвилин. Первинні записи не містили явних часових міток, тому часову послідовність відновлено на підставі порядкового індексу, після чого сформовано синтетичний часовий ряд із кроком 30 хвилин. Цільову змінну визначено як вологість на виході через 30 хвилин, що відповідає зсуву на один крок уперед. До складу вхідних числових ознак включено вміст сухих речовин на вході в барабан, швидкість подачі жому в барабан, температуру на виході з барабана, а за наявності також температуру в камері завантаження, тиск третинного повітря, швидкість обертання барабана та інші релевантні параметри процесу [5; 10]. Пропуски значень, зумовлені формуванням часових ознак на межах ряду, усунуто шляхом вилучення початкових спостережень з неповними вікнами. Поодинокі прогалини усередині часової серії заповнювали прямою та зворотною протяжкою в межах коротких інтервалів, що дає змогу зберегти локальні тренди. Масштабування ознак не виконували, оскільки базова модель коректно працює без їх нормування.

Формування часових ознак. Щоб урахувати інерційність і перехідні режими, до первинних сигналів додано причинно коректні часові ознаки, доступні лише за інформацією на момент t без використання майбутніх значень.

Використано лаги x_{t-L} з $L \in \{1, 2, 3, 4, 6\}$ кроків, що відповідає 0,5, 1, 1,5, 2 і 3 год.

Застосовано ковзні середні MA_w для $w \in \{2, 4, 6\}$ зі згладжуванням на 1, 2 і 3 год. Це стандартний інструмент аналізу часових рядів [11].

Обчислено короткі градієнти $grad_g$ як середню швидкість зміни на вікні з довжиною $g \in \{2, 4, 6\}$ кроків. Зручне наближення має вигляд $grad_g(x)_t \approx (x_t - x_{t-g})/g$.

Такий набір відповідає практиці soft-sensor підходів у промислових сушарках, де інерційні предиктори підсилюють стабільність оцінок у реальному часі [21; 26; 28].

Модель і еталонні методи (базлайни). Базову модель сформовано на ансамблі Random Forest Regressor [2] із такими налаштуваннями: $n_estimators = 400$, $min_samples_leaf = 3$, $random_state = 42$. Вибір RF обґрунтовано стійкістю до шуму, здатністю апроксимувати нелінійні залежності та прийнятною інтерпретованістю за рахунок профілів важливості ознак.

Щоб коректно оцінити приріст якості, застосовано стандартні еталони для часових рядів [11]:

- персистентний прогноз $\hat{y}_{t+1} = y_t$;
- ковзне середнє $MA(3)$ $\hat{y}_{t+1} = (y_t + y_{t-1} + y_{t-2})/3$.

Порівняння з такими базлайнами є критично важливим на квазістаціонарних відрізках, де прості методи нерідко демонструють конкурентні результати.

Оцінювання виконано за схемою перехресної валідації для часових рядів TimeSeriesSplit з двома послідовними фолдами, у яких тестові відрізки хронологічно пізніші за тренувальні [1]. Усі метрики обчислювалися лише на тестових частинах out-of-fold. Горизонт прогнозу становив +30 хв.

Метрики якості. астосовано взаємодоповнювальний набір показників:

- RMSE, MAE – похибки у відсотках вологості.
- R^2 це частка поясненої дисперсії (інтерпретується обережно, якщо робочий діапазон вологості вузький).
- nMAE, nRMSE – значення, нормовані на середню вологість (у відсотках), що полегшує зіставлення результатів між роботами з різними діапазонами.

– Покриття похибок це частка моментів часу, коли $|e| \leq 0,2\%$ та $|e| \leq 0,5\%$. Вона зручна для операційних рішень у СППР.

Додатково визначено важливості ознак за середнім зменшенням нечистоти у Random Forest із подальшим групуванням за фізичним змістом (наприклад, «сухі речовини на вході», «подача жому в барабан», «температура на виході» тощо).

Такий підхід відповідає практиці промислових soft-sensor рішень для сушильних процесів [21; 26].

Обчислення виконано в середовищі Python із використанням загальноприйнятих бібліотек для машинного навчання та аналізу часових рядів; реалізації TimeSeriesSplit і RandomForest взято зі стандартних імплементацій [19]. Усі випадкові компоненти зафіксовано (`random_state = 42`) для забезпечення відтворюваності результатів.

Виклад основного матеріалу.

Кількісні результати. Зведені показники (див. табл. 1) становлять: $RMSE = 0,229\%$, $MAE = 0,180\%$, $R^2 = 0,025$, $nMAE \approx 1,40\%$, $nRMSE \approx 1,78\%$. Поєднання субвідсоткового MAE у фізичних одиницях та горизонту прогнозу +30 хв підтверджує практичну придатність моделі для оперативного контуру, водночас невисоке R^2 зумовлене низькою варіативністю цільового сигналу в квазістаціонарних режимах. Для

прийняття технологічних рішень пріоритетними залишаються $RMSE/MAE$, їх нормовані аналоги $nMAE/nRMSE$ та, за потреби, індикатор time-in-band у реалістично визначеному коридорі якості (див. рис. 1–2).

Нормовані похибки відносно середнього рівня вологості (див. табл. 1) дорівнюють $nMAE \approx 1,40\%$ та $nRMSE \approx 1,74\%$. Це відповідає середній відносній похибці короткострокового прогнозу на рівні близько 1,4% від типового вмісту воло ги в продукті.

Розподіл метрик за фолдами у табл. 1 свідчить про помірну варіабельність, значення між відрізками різняться, $RMSE$ приблизно 0,272 і 0,177, MAE приблизно 0,233 і 0,126, ознак перенавчання не виявлено. Порівняння з наївними еталонами у табл. 2 показує, що на горизонті +30 хв персистентний прогноз y_t та $MA(3)$ мають нижчі $RMSE$ і MAE , що типово для квазістаціонарних ділянок, де y_t уже добре

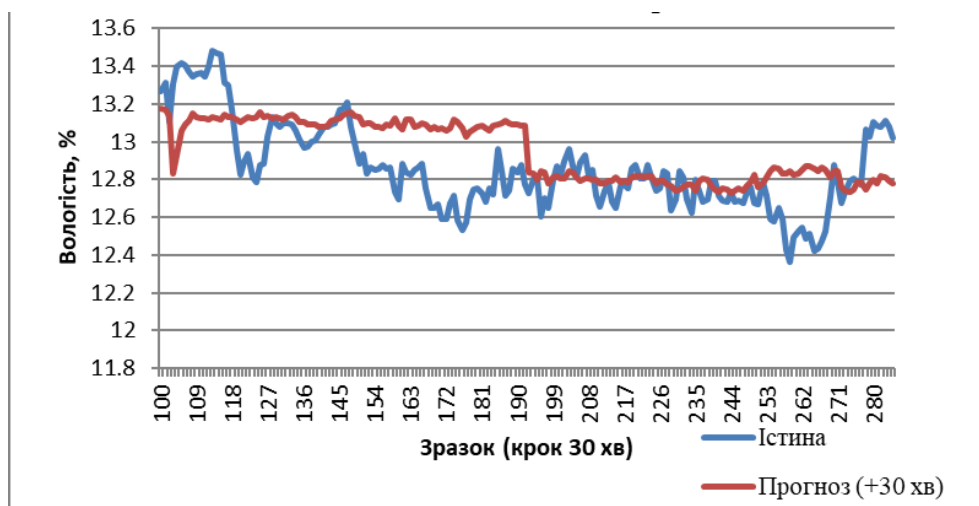


Рис. 1. Вологість: істина vs прогноз

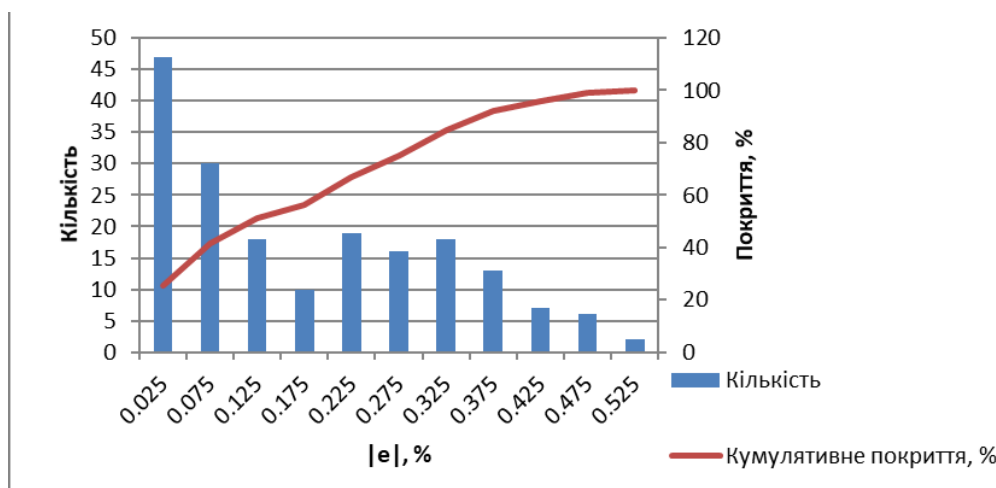


Рис. 2. Абсолютна похибка $|e|$: гістограма

Таблиця 1

Метрики по фолдах

Фолд	RMSE, %	MAE, %	R ²	nMAE, % від середнього	nRMSE, % від середнього	N (тест)
1	0.272	0.233	-0.247	1.80	2.09	93
2	0.177	0.126	-0.294	0.99	1.38	93
Разом для всієї вибірки	0.229	0.180	0.025	1.40	1.78	186

Таблиця 2

Порівняння з базлайнами

Фолд	Модель	RMSE, %	MAE, %	R ²	nMAE, %	nRMSE, %	N (тест)
Фолд 1	Random Forest	0.272	0.233	-0.247	1.798	2.095	93
Фолд 1	Persistence (y_t)	0.081	0.059	0.888	0.455	0.628	93
Фолд 1	MA(3)	0.100	0.077	0.829	0.593	0.775	93
Фолд 2	Random Forest	0.177	0.126	-0.294	0.989	1.385	93
Фолд 2	Persistence (y_t)	0.085	0.067	0.698	0.528	0.669	93
Фолд 2	MA(3)	0.095	0.074	0.629	0.582	0.741	93
Середнє	MA(3)	0.098	0.076	0.729	0.588	0.758	—
Середнє	Persistence (y_t)	0.083	0.063	0.793	0.492	0.648	—
Середнє	Random Forest	0.224	0.180	-0.271	1.394	1.740	—

апроксимує Y_{t+1} . Водночас Random Forest забезпечує високе покриття малих похибок за широким порогом, $|e| \leq 0,5\%$ дорівнює 98,9% для всієї вибірки, для строгого порога $|e| \leq 0,2\%$ отримаємо 56,5% для всієї вибірки, що наведено у табл. 3, модель зберігає інтерпретованість для СППР і має переваги на перехідних режимах. Показник R^2 приблизно 0,025 для всієї вибірки є низьким, з огляду на малу дисперсію цільового сигналу в сталих промислових умовах він менш діагностичний, у виробничому застосуванні доцільно орієнтуватися насамперед на RMSE та MAE, на їх нормовані аналоги nMAE і nRMSE, а за потреби на індикатор time-in-band у коректно визначеному робочому коридорі якості.

Таблиця 3

Покриття похибок Random Forest

Фолд	Покриття $ e \leq 0,2\%$, %	Покриття $ e \leq 0,5\%$, %	N (тест)
Фолд 1	38.7	100.0	93
Фолд 2	74.2	100.0	93
Разом	56.5	98.9	186

Важливість ознак і технологічна інтерпретація. За підсумками оцінки моделі (див. табл. 4) найбільшу індивідуальну вагу має сухі речовини на вході, первинний сигнал 0,1751, на дуже близькому рівні температура на виході, первинний сигнал 0,1662. Сухі речовини на вході, ковзне середнє на 2 кроки 0,1553 і температура на

виході разом ці три предиктори дають приблизно половину сумарної важливості у першій десятці близько 0,497. За групуванням за фізичним змістом сукупність «сухі речовини на вході» та їх похідні первинний сигнал, MA2, lag1, grad2 і grad6 формує близько 39% сумарної важливості. Група «температура на виході» разом із похідними первинний сигнал, MA2, MA4 і lag5 становить близько 26%. Така конфігурація узгоджується з фізикою процесу, матеріальне навантаження визначає обсяг вологи, яку необхідно видалити, а тепловий стан задає доступний потенціал випаровування.

Короткі похідні ознаки помітно допомагають у перехідних режимах. MA2 для сухих речовин на вході має вагу 0,1553, grad2 приблизно 60 хв дає 0,0259, lag1 30 хв дає 0,0184, що підкреслює провідну роль матеріального навантаження. Для температури на виході MA2 має 0,0608, MA4 має 0,0225, lag5 має 0,0136, ці ознаки відбивають теплову інерційність і вплив траєкторії нагріву. Швидкість подачі, первинний сигнал має 0,0136, що логічно, оскільки частину її ефекту вже поглинають ознаки, пов'язані із сухими речовинами та температурою. Параметри з майже сталою варіативністю, зокрема швидкість обертання барабана, очікувано не ввійшли до першої десятки.

У практичному вимірі це означає, що найбільший ефект для керування вологістю на виході дає стабілізація рівня сухих речовин на

вході у поєднанні з обережним регулюванням подачі жому. Такий підхід допомагає оперативно підтримувати потрібний тепловий режим на виході. Використання у СППР коротких ковзних середніх, градієнтів і лагових ознак дає змогу формувати для оператора прогностичні рекомендації під час перехідних режимів, зменшує перерегулювання і скорочує час досягнення цільової вологості.

Таблиця 4

Топ-10 ознак важливості

Ознака	Важливість
Сухі речовини на вході в барабан – первинний сигнал	0.1751
Температура на виході – первинний сигнал	0.1662
Сухі речовини на вході в барабан – ковзне середнє 2 кроки	0.1553
Температура на виході – ковзне середнє 2 кроки	0.0608
Сухі речовини на вході в барабан – градієнт 2 кроки	0.0259
Температура на виході – ковзне середнє 4 кроки	0.0225
Сухі речовини на вході в барабан – запізнення 1 крок	0.0184
Сухі речовини на вході в барабан – градієнт 6 кроків	0.0167
Температура на виході – запізнення 5 кроків	0.0136
Швидкість подачі – первинний сигнал	0.0136

Практичні наслідки для СППР.

У разі прогнозованого зростання вологості першочерговим кроком є стабілізація та, за потреби, помірне зниження рівня сухих речовин на вході в барабан. Це реалізується невеликими змінами швидкості подачі або корекціями попереднього зневоднення. Ефект виявляється протягом одного кроку у 30 хвилин, що підтверджується внеском похідних ознак MA2, lag1 і grad2, див. табл. 4. Швидкість подачі слід змінювати малими інкрементами, щоб уникати різких коливань часу перебування і перерегулювання. На перехідних відрізках, де зростають короткі похідні MA2 і grad2 ключових сигналів, доцільно активувати режим підсилених рекомендацій у СППР.

Температуру на виході доцільно коригувати поступово, що узгоджується з наявністю ознаки lag5 у профілі важливостей, і виконувати переоцінку прогнозу через 30 хвилин. Підвищення температури варто застосовувати лише після стабілізації матеріального навантаження, така послідовність зменшує ризик пересушування та надмірних енерговитрат.

Зіставлення з базовими підходами (див. табл. 2) показує, що на квазістаціонарних відрізках персистентний прогноз і MA3 дають нижчі RMSE і MAE. Отже, у стабільному режимі СППР може обирати простий базлайн. На перехідних ділянках, коли зростають короткі похідні MA2 і grad2 для сухих речовин, подачі та температури, слід покладатися на модель Random Forest, яка точніше відпрацьовує зміни.

Для Random Forest зафіксовано покриття малих похибок, модуль похибки не перевищує 0,2 відсотка у приблизно 56,5 відсотка випадків, і модуль похибки не перевищує 0,5 відсотка у приблизно 98,9 відсотка, див. табл. 3. Доцільно використовувати ці пороги як операційні критерії. Якщо очікувана похибка для поточного сценарію нижча за 0,2 відсотка, рекомендацію можна застосовувати автоматично, якщо вища, потрібне підтвердження оператора.

Порівняння з попередніми дослідженнями.

Для коректного порівняння точності застосовано відносні показники nMAE та nRMSE, оскільки у багатьох роботах розглядають ширші діапазони вологості в лабораторних або перехідних умовах, що механічно підвищує значення R^2 . На нашій лінії за горизонту прогнозу +30 хв отримано nMAE приблизно 1,40 відсотка та nRMSE приблизно 1,78 відсотка. Такий рівень у практиці сушіння вважають достатнім для оперативних корекцій без додаткових запасів [4; 8].

Для роторних і барабанних сушарок процесні дослідження послідовно підкреслюють визначальний вплив матеріального навантаження та температурного режиму сушильного агенту на кінцеву вологість, а також позитивний ефект від упровадження прогнозного керування [5; 10; 16; 18]. Отриманий нами профіль важливостей узгоджується з цими висновками, адже домінують ознаки, пов'язані з сухими речовинами на вході, далі за внеском ідуть швидкість подачі та температура сушильного агенту.

У дослідженнях із застосуванням soft-sensor підходів та ML прогнозів для сушильних процесів показано, що використання агрегованих або інерційних предикторів, зокрема лагів, ковзних середніх і градієнтів, підвищує стабільність прогнозів та їх придатність до оперативного керування [21; 26; 28]. Наші результати узгоджуються з цією тенденцією, оскільки короткі MA, lag і grad суттєво покращують видимість перехідних режимів і безпосередньо формують основу практичних рекомендацій для оператора.

Лабораторні та напівпромислові ML моделі для харчових матеріалів часто демонструють високі значення R^2 на ширших інтервалах вологості

[3; 8; 27]. Це не суперечить отриманим висновкам, адже у стабільних промислових умовах більш інформативними для ухвалення технологічних рішень залишаються MAE і RMSE, їх нормовані аналоги nMAE і nRMSE, а також показник time in band, тоді як саме по собі значення R^2 має обмежену діагностичну цінність.

Висновки. Розроблено модель машинного навчання для прогнозування вологості жому з горизонтом 30 хвилин і подальшого керування насамперед подачею жому. Модель спирається на причинно коректні часові ознаки, зокрема лаги, короткі ковзні середні та градієнти, що відображають інерційність процесу і перехідні режими та забезпечують оперативну придатність прогнозу.

У перехресній валідації для часових рядів з двома фолдами отримано такі зведені метрики для всієї вибірки RMSE 0,229%, MAE 0,180%, R^2 0,025, nMAE приблизно 1,40%, nRMSE приблизно 1,78%. Покриття малих похибок для Random Forest становить модуль похибки не перевищує 0, % у приблизно 56,5% випадків і модуль похибки не перевищує 0,5% у приблизно 98,9%. Такий рівень точності є достатнім для оперативних корекцій без роботи із запасом. На квазістаціонарних ділянках персистентний прогноз і MA3 дають нижчі RMSE і MAE за горизонту 30 хвилин, що мотивує гібридну стратегію. У стабільних режимах доцільні прості базові підходи, у перехідних режимах, коли зростають короткі MA і grad ключових сигналів, доцільно застосовувати ML модель.

Профіль важливостей підтверджує провідну роль групи ознак, пов'язаних із сухими речовинами на вході в барабан. Далі за внеском ідуть температура на виході та швидкість подачі. Короткі похідні типу MA2, lag1 і grad2 найбільше покращують прогноз саме у перехідних режимах. З практичної точки зору пріоритет належить стабілізації матеріального навантаження з подальшими невеликими корекціями швидкості подачі. Підтримка теплового режиму на виході

має бути поступовою з переоцінкою прогнозу через 30 хвилин.

Інтеграція моделі у підсистему підтримки рішень забезпечує проактивні рекомендації оператору з регулярною переоцінкою через 30 хвилин, скорочує перерегулювання, збільшує time in band у реалістично визначеному коридорі якості і не потребує надмірних енерговитрат. У виробничому застосуванні ключовими індикаторами ефективності залишаються RMSE і MAE, їх нормовані аналоги nMAE і nRMSE, а також time in band.

Дані мають синтетичний часовий індекс з кроком 30 хвилин, що ускладнює покадровий аналіз подій у реальному часі. Підсумкові метрики чутливі до вибору ознак і до способу агрегування. На горизонті 30 хвилин у стабільних режимах прості базові підходи часто залишаються дуже сильними, це слід враховувати під час експлуатації.

Перспективи розвитку передбачають упровадження швидкого режиму з горизонтом 15 хвилин для відпрацювання різких переходів, а також перевірку альтернативних алгоритмів, зокрема градієнтного бустингу, послідовних моделей і підходів ансамблювання. Важливим напрямом є організація моніторингу дрейфу з регламентованим перенавчанням моделі. Подальший крок це інтеграція у операторський інтерфейс із відображенням ключових індикаторів time in band та енергоефективності.

Застосування моделі прогнозування вологості на 30 хвилин усуває ключові слабкі місця традиційного керування. Система зменшує запізнення реакції, скорочує потребу у ручному підлаштуванні, дозволяє працювати без запасів, забезпечує субвідсоткову відносну похибку та високе покриття малих помилок і надає операторові інтерпретовані підказки, узгоджені з фізикою процесу сушіння. Це створює основу для промислового впровадження і для подальшої оптимізації енергетичних та якісних показників.

Список літератури:

1. Bergmeir C., Benítez J. M. On the use of cross-validation for time series predictor evaluation. *Information Sciences*. 2012. Vol. 191, №3, P. 192–213. DOI: 10.1016/j.ins.2011.11.028.
2. Breiman L. Random Forests. *Machine Learning*. 2001. Vol. 45, № 1, P. 5–32. DOI: 10.1023/A:1010933404324.
3. Chen F., Gao X., Xia X., Xu J. Using LSTM and PSO techniques for predicting moisture content of poplar fibers by Impulse-cyclone Drying. *PLOS ONE*. 2022. Vol. 17, № 4. DOI: 10.1371/journal.pone.0266186.
4. Dafico L. C. M., Barreira E., Almeida R. M. S. F., Vicente R. Machine learning models applied to moisture assessment in building materials based on surface temperature. *Construction and Building Materials*. 2023. Vol. 405. DOI: 10.1016/j.conbuildmat.2023.133330.
5. Didriksen H. Model-based predictive control of a rotary dryer. *Chemical Engineering Journal*. 2002. Vol. 86, № 1, P. 53–60. DOI:10.1016/S1385-8947(01)00272-8

6. Douglas P. L., Kwade A., Lee P. L., Mallick S. K. Simulation of a rotary dryer for sugar crystalline. *Drying Technology*. 1993. Vol. 11, № 1, P. 129–155. DOI: 10.1080/07373939308916806.
7. Duchesne C., Thibault J., Bazin C. Modelling and dynamic simulation of an industrial rotary dryer. *Developments in Chemical Engineering and Mineral Processing*. 1997. Vol. 5, № 3–4, P. 155–182. DOI: 10.1002/apj.5500050301.
8. Đaković D., Kljajić M., Milivojević N., Doder Đ., Anđelković A. S. Review of energy-related machine learning applications in drying processes. *Energies*. 2024. Vol. 17, № 1, 224. DOI: 10.3390/en17010224.
9. Geankoplis C. J. Transport Processes and Separation Process Principles. 4th ed. Upper Saddle River, NJ: Prentice Hall, 2003, 1026 p.
10. Ghirisan A., Drăgan S., Coța C., Nagy E.-M., Zoltan G., Miclăuș V. Modelling of the sugar beet pulp drying process. *Studia Universitatis Babeș-Bolyai Chemia*. 2020. Vol. 65, № 2, P. 21–28. DOI: 10.24193/subbchem.2020.2.02.
11. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice. 2nd ed. OTexts, 2018. URL: <https://otexts.com/fpp2/>.
12. Iguaz A., Esnoz A., Martínez G., López A., Virseda P. Mathematical modelling and simulation for the drying process of vegetable by-products in a rotary dryer. *Journal of Food Engineering*. 2003. Vol. 59, P. 151–160. DOI: 10.1016/S0260-8774(02)00451-X.
13. Kudra T., Mujumdar A. S. Advanced Drying Technologies. 2nd ed. Boca Raton, FL: CRC Press, 2009.
14. Liu C., Liao C., Peng Y. Microwave sensors and their applications in permittivity measurement. *Sensors*. 2024. Vol. 24, № 23. DOI: 10.3390/s24237696.
15. Miraei Ashtiani S. H. Toward intelligent food drying: Integrating artificial intelligence into drying process. *Drying Technology*. 2024.
16. Mujumdar A. S. (Ed.). Handbook of Industrial Drying. 4th ed. Boca Raton: CRC Press, 2014.
17. Pang S., Jia L., Ding H., Yu Z., Liu Z. Intelligent control in the application of a rotary dryer for reduction of the over-drying of cut tobacco. *Applied Sciences*. 2021. Vol. 11, № 17. DOI: 10.3390/app11178205.
18. Perry R. H., Green D. W., Southard M. Z. (Eds.). Perry's Chemical Engineers' Handbook. 9th ed. New York, NY: McGraw-Hill, 2019.
19. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Duchesnay É. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011. Vol. 12, P. 2825–2830.
20. Prawiranto K., van Mastrigt F., van der Sman R. Physics-based digital twin identifies trade-offs between product quality and drying time. *Frontiers in Sustainable Food Systems*. 2021. Vol. 5. DOI: 10.3389/fsufs.2020.606845.
21. Pla D. L., Kamyar R., Hashemian N., Mehdizadeh H., Moshgbar M. Moisture soft sensor for batch fluid bed dryers: A practical approach. *Powder Technology*. 2018. Vol. 326, P. 69–77. DOI: 10.1016/j.powtec.2017.11.056.
22. Schaßberger J. M. Control-oriented modeling of direct-heat co-current rotary dryers with a focus on demand-side flexibility. *Computers & Chemical Engineering*. 2024.
23. Stöckl G. Employing a multi-resonance microwave sensor for in-line moisture monitoring of fluidized bed agglomeration. *Chemical Engineering Journal*. 2024. Vol. 499. DOI: 10.1016/j.cej.2023.156053.
24. Ullmann's Encyclopedia of Industrial Chemistry. Drying of Solids. Weinheim: Wiley-VCH, 2016.
25. Van der Poel P. W., Schiweck H., Schwartz T. Sugar Technology: Beet and Cane Sugar Manufacture. Berlin: Verlag Dr. Albert Bartens, 1998.
26. Wang Z., Kamyar R., Mehdizadeh H., Pendse P. Y. Moisture soft sensor for agitated pan dryers using a hybrid modeling approach. *International Journal of Pharmaceutics*. 2020. Vol. 586. DOI: 10.1016/j.ijpharm.2020.119518.
27. Yu S., Zheng H., Wilson D. I., Yu W., Young B. R. Integrating image analysis and machine learning for moisture prediction and appearance quality evaluation: A case study of kiwifruit drying pretreatment. *Foods*. 2024. Vol. 13, № 12. DOI: 10.3390/foods13121789.
28. Zambonin E., De Bortoli S., Cancarini D., Beghi A., Tessari M., Susto G. A. Machine learning-based soft sensors for the estimation of laundry moisture content in household dryer appliances. *Energies*. 2019. Vol. 12, № 20. DOI: 10.3390/en12203843.
29. Zhang Y., Fang Z., Li C., Li C. Deep-learning-based model predictive control of an industrial multistage counter-flow paddy drying process. *Foods*. 2023. Vol. 13, № 1. DOI: 10.3390/foods13010043.

Mirkevych O.M. DEVELOPMENT OF A MODEL FOR PREDICTING BEET PULP MOISTURE IN THE DRYING COMPLEX OF SUGAR PRODUCTION

The article examines the development and industrial validation of a machine learning model for forecasting beet pulp moisture in a sugar-production drum-drying complex with a +30-minute horizon. The relevance of the study is driven by the need for proactive control of product quality and energy efficiency in drying, where traditional empirical settings and procedural rules often fail to account for process inertia and transient modes, thereby reducing stability of indicators. The proposed approach relies on production telemetry with a 30-minute sampling interval and causally consistent time features, including lags of 1–6 steps, short moving averages MA(2,4,6), and gradients (2,4,6), which reflect the dynamics of material load and thermal contour. The baseline Random Forest model was evaluated using TimeSeriesSplit with two folds and subsequent aggregation of predictions across the full test set. The achieved accuracy metrics are RMSE 0.229%, MAE 0.180%, R^2 0.025, along with normalized errors $nMAE \approx 1.40\%$ and $nRMSE \approx 1.78\%$. Small-error coverage for Random Forest reaches $|e| \leq 0.2\%$ in $\approx 56.5\%$ of cases and $|e| \leq 0.5\%$ in $\approx 98.9\%$ of observations. Comparison with naïve baselines shows that on quasi-stationary segments the persistence forecast y_t and MA(3) achieve lower RMSE and MAE, whereas in transient regimes the ML model performs better, which justifies a hybrid application policy. The importance profile confirms the dominance of feature groups related to inlet dry matter and outlet temperature, while short derivatives such as MA2, lag1, and grad2 provide sensitivity to rapid changes. The proposed solution generates proactive operator recommendations with re-evaluation every 30 minutes, reduces overshoot, and increases time-in-band within a realistically defined quality corridor, creating a foundation for integration into SCADA or PLC interfaces, deployment of a fast mode with a +15-minute horizon, and implementation of drift monitoring with scheduled model retraining.

Key words: model, beet pulp drying complex, time-series features, moisture forecasting, machine learning, Random Forest.

Дата надходження статті: 19.09.2025

Дата прийняття статті: 07.10.2025

Опубліковано: 16.12.2025