

ИСПОЛЬЗОВАНИЕ СИНТЕТИЧЕСКИХ ДАННЫХ И КО-СИМУЛЯЦИИ ANYLOGIC–PYTHON ДЛЯ РАЗРАБОТКИ GRU-MPC АЛГОРИТМОВ УПРАВЛЕНИЯ МИКРОКЛИМАТОМ ТЕПЛИЦ

Т.В. Костеннов

преподаватель, e-mail: timofey.kostennov@gmail.com

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. В работе рассматривается подход к разработке алгоритмов предиктивного управления микроклиматом тепличных комплексов, основанный на совместном использовании синтетических данных и ко-симуляции AnyLogic–Python. На базе физико-математической модели теплицы формируется расширенный синтетический набор данных, описывающий различные сценарии погодных условий, управляющих воздействий и искажений измерений. На его основе выполняется предварительное обучение рекуррентной нейронной сети с управляемыми зависимостями (GRU) для краткосрочного прогноза параметров микроклимата, после чего модель дообучается на ограниченном объеме реальных производственных данных. Обученный прогнозирующий модуль интегрируется в контур предиктивного управления (GRU-MPC) и исследуется в среде ко-симуляции AnyLogic–Python. Показано, что использование синтетических данных позволяет повысить точность прогноза и устойчивость алгоритма управления в экстремальных режимах по сравнению с обучением исключительно на реальных данных, а ко-симуляция обеспечивает оценку эффективности и безопасности работы контроллера до проведения натурных испытаний.

Ключевые слова: микроклимат теплиц, предиктивное управление, рекуррентные нейронные сети, GRU, синтетические данные, цифровой двойник, ко-симуляция AnyLogic–Python.

Введение

Управление микроклиматом теплиц является одной из ключевых задач ресурсосберегающего растениеводства. Поддержание температуры, влажностного режима (RH/VPD), параметров грунта и концентрации CO_2 непосредственно влияет на урожайность, качество продукции и стрессоустойчивость растений, тогда как отопление, вентиляция, туманообразование и полив формируют значимую долю эксплуатационных энергозатрат [1–3]. Динамика тепличного микроклимата носит нелинейный и нестационарный характер и определяется совокупностью внешних и внутренних факторов, что снижает эффективность простых двухпозиционных регуляторов и затрудняет их настройку в условиях изменчивой и экстремальной погоды [4, 5]. В

связи с этим актуальны предиктивные методы управления, учитывающие динамику объекта, прогноз внешних возмущений, ограничения исполнительных механизмов и требования энергоэффективности [6–8].

Несмотря на результаты, достигнутые рекуррентными нейронными сетями (LSTM, GRU) при прогнозировании временных рядов, их применение в задачах управления теплицами ограничивается дефицитом репрезентативных производственных данных: реальная телеметрия содержит шумы, пропуски и изменения режимов эксплуатации, а экстремальные ситуации представлены недостаточно [9, 10]. Модели, обученные только на таких данных, нередко демонстрируют ухудшение обобщающей способности в непредставленных областях состояний, что повышает риски при переносе алгоритмов на объект.

Решением выступает использование цифровых двойников и синтетических данных [11]. Калиброванная физико-математическая модель теплицы в AnyLogic позволяет безопасно генерировать широкий спектр сценариев, включая экстремальные. На её основе формируется синтетический набор данных для предварительного обучения GRU-прогнозатора микроклимата, после чего модель дообучается на ограниченном реальном наборе данных, что снижает разрыв между синтетическими данными и практикой. Далее GRU-прогнозатор используется в MPC-контуре, реализованном в Python, причём AnyLogic функционирует как «виртуальная теплица» в среде ко-симуляции: на каждом шаге происходит обмен состояниями и управляющими воздействиями [12–14]. Это позволяет тестировать алгоритм в разнообразных условиях, сравнивать настройки MPC и анализировать влияние точности физической модели.

Цель работы – разработать и исследовать единый методический подход к разработке алгоритма предиктивного управления микроклиматом с комбинированным использованием синтетических и реальных данных и последующей проверкой GRU-MPC-контроллера в ко-симуляции AnyLogic–Python. Для этого описывается модель теплицы, разрабатывается метод генерации синтетического набора данных, проектируется архитектура GRU и процедура двухэтапного обучения, формулируется задача MPC с учётом качества регулирования, энергозатрат и ограничений, реализуется среда ко-симуляции и проводится сравнение предложенного подхода с базовыми регуляторами. Проверяемая гипотеза состоит в том, что совместное использование синтетических данных, реальных данных и ко-симуляции повышает точность прогноза, устойчивость и эффективность управления и снижает риски натурных испытаний по сравнению с подходом, основанным только на реальных данных [15].

1. Математическая модель теплицы

1.1. Структура объекта управления и основные переменные

Рассматривается отдельный тепличный модуль (одна или несколько одинаковых секций), описываемый сосредоточенной моделью с усреднёнными по объёму параметрами. Внутренний объём теплицы находится в тепло- и влагомассообмене с грунтом, прозрачным покрытием и внешней средой, а также с системами отопле-

ния, вентиляции и туманообразования [3, 4].

Основные *состояния* модели:

$$x = \begin{bmatrix} T_a & w_a & T_g & T_c \end{bmatrix}^T,$$

где T_a – температура воздуха, w_a – массовая концентрация водяного пара, T_g – эффективная температура верхнего слоя грунта, T_c – эффективная температура ограждений. При необходимости модель может быть расширена концентрацией C_{CO_2} .

Управляющие *воздействия*:

$$u = \begin{bmatrix} u_h & u_v & u_f & u_s \end{bmatrix}^T,$$

где u_h – относительная мощность отопления, u_v – степень вентиляции, u_f – интенсивность туманообразования/орошения, u_s – положение штор. При наличии досветки вводится u_l как дополнительное управление или возмущение.

Внешние *возмущения*:

$$d = \begin{bmatrix} T_{out} & w_{out} & G_{sol} & v_{wind} \end{bmatrix}^T,$$

где T_{out} , w_{out} – параметры наружного воздуха, G_{sol} – суммарная солнечная радиация, v_{wind} – скорость ветра.

1.2. Физико-математическая модель тепло- и влагопереноса

Модель задаётся системой ОДУ для балансов энергии и влаги в характерных подсистемах.

Баланс энергии для воздуха теплицы

Температура воздуха T_a описывается уравнением

$$C_a \frac{dT_a}{dt} = Q_{sol} + Q_{heat} + Q_{soil} + Q_{cover} + Q_{vent} + Q_{lat} + Q_{int}, \quad (1)$$

где $C_a = \rho_a c_{p,a} V_a$ – эффективная теплоёмкость воздушного объёма.

Основные слагаемые:

$$Q_{sol} = \tau_s(G_{sol}, u_s) G_{sol} A_c, \quad (2)$$

$$Q_{heat} = \eta_h F_h^{\max} u_h, \quad (3)$$

$$Q_{soil} = K_{ag}(T_g - T_a), \quad Q_{cover} = K_{ac}(T_c - T_a), \quad (4)$$

$$Q_{vent} = \rho_a c_{p,a} \dot{V}_{vent}(u_v, v_{wind})(T_{out} - T_a), \quad (5)$$

$$Q_{lat} = -L_v \dot{m}_{evap}, \quad (6)$$

где $\dot{V}_{vent}$ – расход вентиляционного воздуха, $\dot{m}_{evap}$ – массовая скорость испарения влаги. Q_{int} агрегирует внутренние тепловыделения.

Баланс влаги в воздухе

Влагосодержание воздуха w_a изменяется под действием испарения/транспирации, туманообразования, вентиляции и конденсации:

$$V_a \frac{dw_a}{dt} = \dot{m}_{\text{evap}}(T_a, w_a, \text{VPD}) + \dot{m}_{\text{fog}}(u_f) + \dot{m}_{\text{irr}} - \dot{m}_{\text{vent}}(u_v, v_{\text{wind}}) - \dot{m}_{\text{cond}}(T_c, w_a), \quad (7)$$

где потоки влаги задаются полуэмпирическими зависимостями, например

$$\dot{m}_{\text{evap}} = k_{\text{evap}} \text{VPD}(T_a, w_a) A_{\text{leaf}}. \quad (8)$$

Поток через вентиляцию:

$$\dot{m}_{\text{vent}} = \dot{V}_{\text{vent}}(u_v, v_{\text{wind}}) (w_a - w_{\text{out}}). \quad (9)$$

$\dot{m}_{\text{cond}}$ описывает конденсацию на внутренней поверхности покрытия.

Связь w_a с RH и VPD задаётся стандартными термодинамическими соотношениями:

$$\text{VPD} = e_{\text{sat}}(T_a) - e_v(T_a, w_a), \quad (10)$$

где e_{sat} – давление насыщенного пара, e_v – фактическое парциальное давление водяного пара.

Модель грунта и ограждений

Температура грунта T_g :

$$C_g \frac{dT_g}{dt} = K_{ag}(T_a - T_g) + Q_{g,\text{solar}} - Q_{g,\text{loss}}, \quad (11)$$

где $Q_{g,\text{solar}}$ – поглощённая радиация, $Q_{g,\text{loss}}$ – теплоотвод в грунт.

Температура ограждений T_c :

$$C_c \frac{dT_c}{dt} = K_{ac}(T_a - T_c) + K_{c,\text{out}}(T_{\text{out}} - T_c) + Q_{c,\text{rad}}, \quad (12)$$

где $K_{c,\text{out}}$ учитывает теплообмен с наружным воздухом и ветром, $Q_{c,\text{rad}}$ суммирует радиационный обмен.

Система (1)–(12) определяет динамику $x(t)$ при заданных $u(t)$ и $d(t)$. Численное интегрирование с шагом 5–15 минут реализовано в AnyLogic.

1.3. Учитываемые возмущения и управляющие воздействия

Модель чётко разделяет *управления* u и *возмущения* d , что важно для генерации синтетических данных и постановки МРС.

Управляющие воздействия $u_h, u_v, u_f, u_s \in [0, 1]$ описывают соответственно относительную мощность отопления, вентиляции, туманообразование/орошение и положение штор. При необходимости вводится управление досветкой u_l . Для оборудования учитываются ограничения по диапазону, скорости изменения и минимальному времени включения.

Возмущения $T_{\text{out}}(t)$, $w_{\text{out}}(t)$, $G_{\text{sol}}(t)$, $v_{\text{wind}}(t)$ задаются по данным метеостанции или с помощью синтетически сгенерированных профилей, в том числе с резкими изменениями для формирования экстремальных сценариев.

1.4. Кратко о калибровке и проверке модели

Параметры модели θ (коэффициенты тепло- и влагопереноса, эффективные теплоёмкости, параметры оборудования и др.) оцениваются по реальным данным через минимизацию невязки между моделью и измерениями:

$$J(\theta) = \sum_k \sum_{i \in \mathcal{Y}} w_i (y_i^{\text{mod}}(t_k, \theta) - y_i^{\text{meas}}(t_k))^2 \rightarrow \min_{\theta \in \Theta}, \quad (13)$$

где $\mathcal{Y}$ – набор наблюдаемых выходов (прежде всего T_a и RH/VPD), Θ – физически допустимая область параметров.

Используются стандартные комбинации градиентных и стохастических методов; детали не приводятся. После идентификации модель проверяется на независимых данных по RMSE, отсутствию выраженного смещения и корректному воспроизведению реакций на изменения погоды и управления. При достижении приемлемой точности модель рассматривается как цифровой двойник, достаточный для генерации синтетических данных и ко-симуляции AnyLogic–Python [5, 14].

2. Генерация синтетических данных

2.1. Сценарии погодных условий и внешних воздействий

Синтетический набор данных формируется путём прогонов цифрового двойника под набором погодных сценариев, т. е. временных профилей

$$T_{\text{out}}(t), w_{\text{out}}(t), G_{\text{sol}}(t), v_{\text{wind}}(t)$$

длительностью от нескольких суток до нескольких недель с шагом дискретизации 5–15 минут [1, 11].

Формируются несколько классов сценариев:

- штатные сезонные – типичные весенне-летние и осенние условия;
- холодные экстремальные – низкие температуры, быстрые похолодания;
- жаркие экстремальные – высокие температуры и радиация;
- переходные – быстрые изменения облачности, ветра и температуры.

Часть сценариев основана на реальных метеоданных со сглаживанием и масштабированием, часть – синтетическая (гармонические компоненты + шум). Контролируются диапазоны $T_{\text{out}}, w_{\text{out}}, G_{\text{sol}}, v_{\text{wind}}$, суточные циклы и базовые статистики.

2.2. Сценарии управления

Для каждого погодного профиля задаётся несколько вариантов управления, чтобы получить разнообразие траекторий:

- релейные регуляторы с гистерезисом по T_a и RH/VPD с разными порогами;
- простые пропорциональные/псевдо-PID-схемы по отклонению от уставок;
- сценарии с программным расписанием уставок (день/ночь, закалка и т.п.);
- возбуждающие (идентификационные) воздействия: ступенчатые и псевдослучайные изменения u_h, u_v, u_f в допустимых пределах.

Во всех случаях соблюдаются физические и агротехнические ограничения: $u_h, u_v, u_f, u_s \in [0, 1]$, учитываются минимальное время работы оборудования и недопустимость длительных экстремальных условий для растений.

2.3. Моделирование шумов, задержек и пропусков измерений

Чтобы синтетические данные были сопоставимы с реальными, к «идеальным» выходам модели добавляются искажения, имитирующие измерения:

- аддитивный шум с параметрами, соответствующими датчикам (температура, RH/VPD, T_g);
- медленный дрейф калибровки для части эпизодов;
- квантование (ограниченное разрешение измерений);
- задержка между фактическим состоянием и доступными измерениями;
- редкие пропуски и выбросы, обрабатываемые теми же методами фильтрации, что и для реальных данных.

После добавления шумов и пропусков к данным применяется единая технологическая процедура предобработки, аналогичный реальному (фильтрация, интерполяция, отсечение явных артефактов).

2.4. Структура синтетического набора данных и разделение train/val/test

На каждом шаге t_k сохраняются:

- текущие измеряемые переменные микроклимата (T_a , RH/VPD, при необходимости T_g и др.);
- управляющие воздействия $u_h(t_k)$, $u_v(t_k)$, $u_f(t_k)$, $u_s(t_k)$;
- внешние возмущения $T_{out}(t_k)$, $w_{out}(t_k)$, $G_{sol}(t_k)$, $v_{wind}(t_k)$;
- значения микроклимата на горизонте прогноза (1–3 часа вперёд).

Данные организованы по принципу «сценарий $\rightarrow$ временной ряд»; объём синтетического набора составляет десятки тысяч временных шагов (несколько виртуальных сезонов). Деление на train/val/test выполняется *по сценариям*, что исключает утечку информации:

- train – основная часть сценариев всех типов;
- val – отдельные сценарии (в т.ч. экстремальные) для подбора гиперпараметров;
- test – сценарии, полностью исключённые из обучения и валидации.

На основе этих временных рядов формируются скользящие окна для обучения GRU (2–6 часов истории и 1–3 часа прогноза).

2.5. Оценка реалистичности: сравнение синтетических и реальных данных

Реалистичность синтетического набора данных проверяется сравнением с историческими данными теплицы (1–2 сезона):

- сравниваются диапазоны и распределения T_a , RH/VPD, управляющих воздействий и внешних возмущений;
- анализируются усреднённые суточные профили и сезонные изменения;

- рассматриваются распределения приращений, автокорреляции и взаимные корреляции (например, между G_{sol} и ростом T_a);
- сравниваются амплитуда и характер шумов, частота выбросов и пропусков.

При необходимости корректируются параметры модели и генератора сценариев. Цель – не точное совпадение распределений, а физически правдоподобные и динамически корректные данные, расширяющие область сценариев за счёт экстремальных режимов.

3. Обучение GRU-модели на синтетических и реальных данных

3.1. Архитектура GRU-прогнозатора и формирование входных/выходных векторов

GRU-прогнозатор используется как динамическая модель микроклимата, связывающая историю состояний, управления и внешних возмущений с будущей траекторией. В дискретном времени $t_k = k\Delta t$:

$$y_k = \begin{bmatrix} T_a(k) \\ \text{VPD}(k) \end{bmatrix}, \quad u_k = \begin{bmatrix} u_h(k) \\ u_v(k) \\ u_f(k) \\ u_s(k) \end{bmatrix}, \quad d_k = \begin{bmatrix} T_{\text{out}}(k) \\ w_{\text{out}}(k) \\ G_{\text{sol}}(k) \\ v_{\text{wind}}(k) \end{bmatrix},$$

$$z_k = \begin{bmatrix} y_k \\ u_k \\ d_k \end{bmatrix}.$$

Для обучения строятся окна длиной L :

$$Z_k = \{z_{k-L+1}, \dots, z_k\}, \quad Y_k^{(\text{tar})} = \{y_{k+1}, \dots, y_{k+H}\},$$

где H – горизонт прогноза.

Используется один–два слоя GRU (32–128 скрытых нейронов) и полносвязный выходной слой (или маленькая MLP-«голова»), формирующий вектор

$$\hat{Y}_k = [\hat{y}_{k+1}, \dots, \hat{y}_{k+H}].$$

Входы стандартизируются (вычитание среднего и деление на стандартное отклонение). Модель реализует отображение

$$\hat{Y}_k = f_{\theta}(Z_k),$$

где θ – параметры сети. Обучение формулируется как задача многовыходной регрессии [9, 10].

3.2. Этап предварительного обучения (pre-training) на синтетическом наборе

На первом этапе GRU обучается на синтетическом наборе данных

$$\mathcal{D}_{\text{syn}} = \{(Z_k^{(\text{syn})}, Y_k^{(\text{syn})})\}$$

со случайной инициализацией параметров и разделением на train/val по сценариям. Используется Adam с шагом порядка 10^{-3} и L2-регуляризацией:

$$\mathcal{L}_{\text{syn}}(\theta) = \frac{1}{N_{\text{syn}}} \sum_k \sum_{h=1}^H \sum_j w_{j,h} \left(y_{k+h,j}^{(\text{syn})} - \hat{y}_{k+h,j} \right)^2 + \lambda \|\theta\|_2^2.$$

Качество контролируется по RMSE/MAE и R^2 для T_a и VPD, а остановка – по критерию ранней остановки на валидации. На этом этапе сеть выучивает базовую динамику теплицы, суточные и сезонные паттерны и реакции на экстремальные воздействия. Полученные параметры θ_{pre} используются как старт для дообучения.

3.3. Этап дообучения (fine-tuning) на ограниченном реальном наборе данных

На втором этапе GRU адаптируется к реальному объекту по набору данных

$$\mathcal{D}_{\text{real}} = \{(Z_k^{(\text{real})}, Y_k^{(\text{real})})\},$$

полученному из 1–2 сезонов реальной эксплуатации после предобработки. Архитектура и структура признаков сохраняются, параметры инициализируются из θ_{pre} , шаг обучения уменьшается, при необходимости часть слоёв «замораживается». Целевая функция:

$$\mathcal{L}_{\text{real}}(\theta) = \frac{1}{N_{\text{real}}} \sum_k \sum_{h=1}^H \sum_j w_{j,h} \left(y_{k+h,j}^{(\text{real})} - \hat{y}_{k+h,j} \right)^2 + \lambda_{\text{real}} \|\theta\|_2^2.$$

Используются те же метрики (RMSE, MAE, R^2), дополнительно анализируются смещение прогноза и ошибки в экстремальных режимах. Объём $\mathcal{D}_{\text{real}}$ существенно меньше объёма $\mathcal{D}_{\text{syn}}$, поэтому предварительное обучение играет роль структурной регуляризации.

3.4. Сравнение с моделью, обученной только на реальных данных

Для оценки влияния синтетических данных используется базовая модель, обученная только на $\mathcal{D}_{\text{real}}$ с той же архитектурой, признаками и процедурами нормировки. Обучение ведётся с нуля, с подбором гиперпараметров по валидации.

Сравнение двух моделей (**GRU (real-only)** и **GRU (syn+real)**) проводится на независимом реальном тестовом наборе по RMSE/MAE, R^2 , распределениям ошибок и поведению в экстремальных сценариях. Модель **GRU (syn+real)** показывает более низкие ошибки и более устойчивое качество прогнозирования, что подтверждает пользу предварительного обучения на синтетических данных и согласуется с результатами современных работ по предиктивному управлению микроклиматом теплиц [5, 15].

4. GRU-MPC и ко-симуляция AnyLogic–Python

4.1. Постановка задачи MPC и целевая функция

MPC формулируется с конечным горизонтом N и принципом скользящего горизонта. Управляемые переменные – температура воздуха T_a и показатель влажностного режима $z_{\text{hum}} = \text{VPD}$ с допустимыми диапазонами $[T_a^{\min}, T_a^{\max}]$, $[z_{\text{hum}}^{\min}, z_{\text{hum}}^{\max}]$.

Целевая функция:

$$J = \sum_{h=1}^N (\|T_{a,k+h} - T_{a,k+h}^{\text{ref}}\|_{q_T}^2 + \|z_{\text{hum},k+h} - z_{\text{hum},k+h}^{\text{ref}}\|_{q_H}^2) + \sum_{h=0}^{N-1} (\|u_{k+h}\|_{r_u}^2 + \|\Delta u_{k+h}\|_{r_{\Delta u}}^2), \quad (14)$$

где $u_{k+h} = [u_h, u_v, u_f, u_s]^T$, $\Delta u_{k+h} = u_{k+h} - u_{k+h-1}$. Первые слагаемые отвечают за качество микроклимата, вторые – за энергозатраты и плавность управления.

Ограничения по состоянию и управляющим воздействиям:

$$T_a^{\min} \leq T_{a,k+h} \leq T_a^{\max}, \quad z_{\text{hum}}^{\min} \leq z_{\text{hum},k+h} \leq z_{\text{hum}}^{\max}, \quad (15)$$

$$u^{\min} \leq u_{k+h} \leq u^{\max}, \quad \Delta u^{\min} \leq \Delta u_{k+h} \leq \Delta u^{\max}. \quad (16)$$

Внешние возмущения

$$d_{k+h} = [T_{\text{out}}, w_{\text{out}}, G_{\text{sol}}, v_{\text{wind}}]^T$$

на горизонте считаются известными (реальный прогноз или заранее заданный сценарий). На каждом шаге реализуется только первый элемент оптимальной траектории u_k^* , далее задача решается заново [6–8].

4.2. Интеграция GRU-прогнозов в контур MPC

Динамика теплицы в MPC аппроксимируется GRU-моделью. На шаге k контроллер имеет:

- текущий вектор y_k и последнее управление u_{k-1} ;
- прогноз возмущений $\{d_k, \dots, d_{k+N}\}$;
- историческое окно $\{z_{k-L+1}, \dots, z_k\}$.

Переменной оптимизации служит траектория $\mathbf{u} = (u_k, \dots, u_{k+N-1})$. Для заданной $\mathbf{u}$ GRU итеративно строит прогноз:

$$\hat{y}_{k+1} = f_{\theta}(Z_k, u_k, d_{k+1}), \quad (17)$$

далее окно сдвигается и получаются $\hat{y}_{k+2}, \dots, \hat{y}_{k+N}$. Это соответствует развёртке GRU по горизонту MPC с общими весами.

Поскольку f_{θ} дифференцируема по входам, задача MPC решается как нелинейная оптимизация с использованием автоматического дифференцирования: градиенты $\nabla_{\mathbf{u}} J$ вычисляются через backpropagation, оптимизация ведётся градиентными или квазиньютоновскими методами с учётом ограничений, по аналогии с работами по интеграции нейросетевых моделей в контуры предиктивного управления микроклиматом теплиц [5, 13, 14].

Для практической реализуемости выбирается умеренный горизонт N (например, 12–24 шагов), ограниченное число итераций оптимизатора и разумное начальное приближение (решение с предыдущего шага или базовая стратегия).

4.3. Архитектура ко-симуляции AnyLogic–Python и протокол обмена

Ко-симуляция AnyLogic–Python позволяет тестировать GRU-MPC на цифровом двойнике до натуральных испытаний. AnyLogic реализует модель теплицы и погодные сценарии, Python – GRU-прогнозатор и MPC.

При обмене с шагом Δt выполняются:

1. AnyLogic на шаге k вычисляет состояние x_k , формирует измерения y_k и возмущения d_k .
2. Через интерфейс (например, JSON по локальному TCP/HTTP) AnyLogic передаёт в Python пакет данных $\{y_k, u_{k-1}, d_k, \text{историческое окно}\}$.
3. Python обновляет буфер окна, при необходимости нормирует данные, решает задачу MPC и возвращает u_k^* .
4. AnyLogic применяет u_k^* на интервале $[t_k, t_{k+1})$ и интегрирует ОДУ до t_{k+1} .

Далее $k \leftarrow k + 1$ и цикл повторяется. Оба компонента работают синхронно: AnyLogic ждёт управления, Python – новых данных. Такая схема обеспечивает согласование «виртуального времени», удобный запуск серий сценариев и логирование всех переменных для анализа, и может быть реализована, в частности, с использованием библиотеки Alpyne для связи AnyLogic и Python [11, 12, 16].

4.4. Набор сценариев ко-симуляции (штатные и экстремальные режимы)

Для оценки эффективности и устойчивости GRU-MPC используется набор сценариев (см. раздел 2.), включающий:

- штатные суточные циклы (типичные дни с умеренными колебаниями);
- холодные экстремальные сценарии (длительные похолодания, ночные минимумы);
- жаркие сценарии (волны жары с высокой радиацией);
- быстрые переходы (резкая смена облачности, скачки температуры и ветра);
- комбинированные стресс-сценарии (неблагоприятная погода + ограничения по оборудованию).

Каждый сценарий проигрывается с GRU-MPC и с базовой стратегией (релейным или гистерезисным регулятором). Для каждого прогона рассчитываются интегральные ошибки (IAE/ISE по T_a и z_{hum}), суммарное энергопотребление, число переключений оборудования и показатели устойчивости (выходы за пределы, наличие/отсутствие колебаний). Это служит основой для сравнительного анализа в разделе 5..

5. Экспериментальные результаты

5.1. Качество прогноза: синтетические данные и реальные данные против только реальных данных

Качество GRU-прогнозатора оценивалось на независимой тестовой выборке реальных данных. Сравнивались две модели:

- **GRU (real-only)** – обучение только на реальном наборе данных;
- **GRU (syn+real)** – предварительное обучение на синтетических данных и дообучение на реальных данных.

Архитектура, схема признаков, нормировка и метрики совпадали. Оценка проводилась по RMSE и MAE для T_a и z_{hum} на горизонтах 1, 2 и 3 часа, коэффициенту R^2 и распределениям ошибок.

Модель **GRU (syn+real)** показывает более низкие RMSE/MAE и более высокий R^2 , особенно для больших горизонтов и в динамичных режимах (быстрая смена погоды, переходы день/ночь). В штатных условиях качество моделей сопоставимо, тогда как в экстремальных ситуациях преимущество синтетически предобученной модели выражено сильнее, в том числе за счёт уменьшения доли крупных ошибок. Смещение (bias) по температуре и VPD для **GRU (syn+real)** ближе к нулю, что важно для MPC.

5.2. Качество управления в ко-симуляции: GRU-MPC vs базовая стратегия

Каждый тестовый сценарий ко-симуляции проигрывается в двух вариантах:

- **GRU-MPC** – MPC с нейросетевым предсказателем **GRU (syn+real)**;
- **Базовая стратегия** – релейные или гистерезисные регуляторы с теми же диапазонами T_a и z_{hum} .

Для каждого варианта регистрируются траектории микроклимата, управления и возмущений. Оценка проводится по:

- интегральным ошибкам

$$\text{IAE}_T = \sum_k |T_a(k) - T_a^{\text{ref}}(k)|, \quad \text{ISE}_T = \sum_k (T_a(k) - T_a^{\text{ref}}(k))^2,$$

и аналогичным показателям по z_{hum} ;

- суммарному энергопотреблению

$$E = \sum_k (\alpha_h u_h(k) + \alpha_v u_v(k)) \Delta t;$$

- числу переключений основных исполнительных механизмов и распределению длительностей циклов работы.

GRU-MPC обеспечивает более точное поддержание T_a и z_{hum} в заданных диапазонах (меньше IAE/ISE) и более плавное управление с меньшим числом переключений. В большинстве сценариев наблюдается снижение энергозатрат, особенно при

нестабильной погоде, благодаря упреждающим действиям. В экстремальных сценариях GRU-MPC лучше удерживает микроклимат в допустимых границах при ограниченных ресурсах (мощность отопления, максимальная вентиляция). В отдельных случаях MPC может расходовать немного больше энергии, существенно улучшая качество микроклимата; такие компромиссы зависят от выбранных весов в функционале. Полученные эффекты находятся в одном ряду с результатами других работ по предиктивному управлению микроклиматом теплиц [5, 13, 15].

5.3. Анализ устойчивости и энергетической эффективности

Устойчивость и энергоэффективность анализировались по результатам длительных прогонов:

- переходы между режимами (день/ночь, смена уставок) сопровождаются сглаженными траекториями без значимых перерегулирований;
- при резких изменениях внешних условий контроллер реагирует упреждающе, уменьшая амплитуду отклонений микроклимата;
- на длительных интервалах не наблюдается самоподдерживающихся колебаний; система возвращается в допустимую область после сильных возмущений.

С точки зрения энергопотребления GRU-MPC:

- уменьшает число «чрезмерных» включений оборудования на полную мощность;
- распределяет действия по горизонту, позволяя работать ближе к нижним границам допустимых диапазонов в периоды высокого спроса на энергию;
- рационально комбинирует отопление, вентиляцию, шторы и туманообразование.

Эффект по энергосбережению зависит от веса энергетического слагаемого в целевой функции и конкретных агротехнических приоритетов.

5.4. Обсуждение ограничений и чувствительности к параметрам модели

Качество GRU-MPC ограничивается:

- точностью цифрового двойника: систематические ошибки модели (например, в оценке теплотерь или эффективности вентиляции) частично переносятся в синтетический набор данных и нейросетевой предсказатель; дообучение на реальных данных уменьшает, но не устраняет этот разрыв;
- выбором и охватом сценариев: недостаточное разнообразие ухудшает поведение в редких режимах, чрезмерное насыщение малореалистичными экстремумами может ухудшить качество в типичных условиях;
- настройкой гиперпараметров GRU и MPC (длина окна, горизонт, веса, ограничения по Δu), влияющих на баланс между упреждением, устойчивостью и вычислительной сложностью;
- качеством прогноза внешних возмущений: в ко-симуляции возмущения известны точно, в реальности – с погрешностью метеопрогноза, что требует робастных/стохастических модификаций MPC;

- доступными вычислительными ресурсами: решение MPC-задачи с нейросетевой динамикой в реальном времени требует оптимизированной реализации и, при необходимости, специализированного оборудования.

Тем не менее результаты ко-симуляции подтверждают, что использование синтетических данных и ко-симуляции AnyLogic–Python позволяет ускорить и безопасно провести разработку предиктивных алгоритмов управления микроклиматом теплиц и обеспечить выигрыш по качеству регулирования и энергоэффективности по сравнению с базовыми стратегиями.

Заключение

Разработана и исследована сквозная методика построения алгоритма предиктивного управления микроклиматом теплицы (GRU-MPC), основанная на совместном использовании синтетических и реальных данных и их интеграции в среду ко-симуляции AnyLogic–Python. Калиброванный цифровой двойник применён для генерации синтетической выборки по набору штатных и экстремальных сценариев погоды и управления, включая имитацию шумов и пропусков измерений. Предварительное обучение GRU-прогнозатора на синтетической выборке с последующим дообучением на реальных данных позволило снизить расхождение между моделью и объектом и повысить устойчивость прогноза в режимах, слабо представленных в реальной телеметрии. В ко-симуляционных экспериментах показано, что GRU-MPC по сравнению с базовыми двухпозиционными и гистерезисными регуляторами обеспечивает более точное удержание микроклимата в допустимых диапазонах, уменьшает интегральные ошибки регулирования и в большинстве сценариев снижает энергозатраты при одновременном сокращении числа переключений исполнительных механизмов.

Литература

1. Chen S. A Review of Environmental Control Strategies and Models for Modern Agricultural Greenhouses // *Sensors*. 2025. V. 25. P. 1388.
2. Боярцева В.К. Микроклимат теплиц. Справочник садовода. 2010. 420 с.
3. Байрамов Р.Б., Рыбакова Л.Е. Микроклимат теплиц на солнечном обогреве. Ашхабад: Ылым, 1983. 84 с.
4. Vanthoor B.H.E., Stanghellini C., van Henten E.J., de Visser P.H.B. A methodology for model-based greenhouse design: Part 1. A greenhouse climate model for a broad range of designs and climates // *Biosystems Engineering*. 2011. V. 110, No. 4. P. 363–377.
5. Mahmood F., Govindan R., Bermak A., Yang D., Al-Ansari T. Data-driven robust model predictive control for greenhouse temperature control and energy utilisation assessment // *Applied Energy*. 2023. V. 343. Art. 121190.
6. Rawlings J.B., Mayne D.Q., Diehl M. *Model Predictive Control: Theory, Computation, and Design*. Madison, WI: Nob Hill Publishing, 2017. 623 p.
7. Camacho E.F., Bordons C. *Model Predictive Control*. London: Springer-Verlag, 2004. 405 p.
8. Qin S.J., Badgwell T.A. A survey of industrial model predictive control technology // *Control Engineering Practice*. 2003. V. 11, No. 7. P. 733–764.

9. Liu Y., Li D., Wan S., Wang F., Dou W., Xu X., et al. A long short-term memory-based model for greenhouse climate prediction // International Journal of Intelligent Systems. 2022. V. 37, No. 1. P. 135–151.
10. Zhang G. et al. Predicting greenhouse air temperature using LSTM-AT // Transactions of the Chinese Society of Agricultural Engineering. 2024. V. 40, No. 18. P. 194–201.
11. Ariesen-Verschuur N., Verdouw C., Tekinerdogan B. Digital Twins in greenhouse horticulture: A review // Computers and Electronics in Agriculture. 2022. V. 199. Art. 107183.
12. McClellan A., Lorenzetti J., Pavone M., Farhat C. A physics-based digital twin for model predictive control of autonomous unmanned aerial vehicle landing // Philosophical Transactions of the Royal Society A. 2022. V. 380. Art. 20210204.
13. Fitz-Rodríguez E., Kacira M., Villarreal-Guerrero F., Giacomelli G.A., Linker R., Kubota C., Arbel A. Neural network predictive control in a naturally ventilated and fog cooled greenhouse // Acta Horticulturae. 2012. V. 952. P. 45–52.
14. Chen W.H. Data-driven robust optimization for greenhouse temperature and CO_2 concentration control // Chemical Engineering Transactions. 2020. V. 81. P. 671–676.
15. Soumo E.A., Calisti M., Polydoros A. Data-Driven Greenhouse Climate Regulation in Lettuce Cultivation Using BiLSTM and GRU Predictive Control. Preprint, 2025. URL: <https://arxiv.org/abs/2507.21669> (дата обращения: 16.12.2025).
16. Alpyne: A Python library for interactively running AnyLogic models. URL: <https://github.com/the-anylogic-company/Alpyne> (дата обращения: 16.12.2025).

SYNTHETIC DATA AND ANYLOGIC–PYTHON CO-SIMULATION FOR DEVELOPMENT OF GRU-MPC GREENHOUSE CLIMATE CONTROLLERS

T.V. Kostenov

Instructor, e-mail: timofey.kostenov@gmail.com

Dostoevsky Omsk State University, Omsk, Russia

Abstract. This paper presents an approach to the development of model predictive control (MPC) algorithms for greenhouse climate management based on synthetic data generation and AnyLogic–Python co-simulation. A physics-based greenhouse model is used to generate an extended synthetic dataset covering diverse weather conditions, control strategies and measurement distortions. This dataset is used to pre-train a gated recurrent unit (GRU) neural network for short-term climate prediction, which is then fine-tuned on a limited amount of real operational data. The trained forecasting module is integrated into a GRU-MPC controller and evaluated within an AnyLogic–Python co-simulation environment. The results demonstrate that combining synthetic and real data improves prediction accuracy and control robustness in extreme operating regimes compared to training on real-world data only, while co-simulation enables assessment of the controller's performance and safety prior to field deployment.

Keywords: greenhouse climate, model predictive control, recurrent neural networks, GRU, synthetic data, digital twin, AnyLogic–Python co-simulation.

Дата поступления в редакцию: 02.12.2025