

Modeling Greenhouse Cucumber Evapotranspiration Using Machine Learning: A Random Forest Approach Versus Traditional and Non-linear Crop Coefficients

Morteza Khoshsimaie Chenar^{1*}, Arash Tafteh², Niazali Ebrahimipak³

¹ PhD Student in Irrigation and Drainage Engineering, Department of Irrigation and Reclamation Engineering, Faculty of Agricultural, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

² Associate Professor of Soil and Water Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran.

³ Associate Professor of Soil and Water Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran.

Extended Abstract

Introduction

Accurate estimation of crop evapotranspiration (ET_c) is fundamental for the development of efficient irrigation strategies in greenhouse systems, where environmental conditions differ significantly from open-field farming. In Iran, greenhouse agriculture, particularly for cucumbers (*Cucumis sativus* L.), has expanded considerably, making irrigation optimization a critical priority. The specific microclimate within greenhouses, including controlled humidity, temperature, and solar radiation levels, affects plant water needs, requiring tailored approaches to predict ET_c . Traditional models, like the FAO 56 crop coefficient (K_c) method, provide a standardized way to estimate ET_c but are generally suited to field crops under variable outdoor conditions. The limitations of fixed K_c values in capturing the complexity of greenhouse environments have prompted the exploration of alternative models. In recent years, machine learning (ML) techniques, especially ensemble methods like the Random Forest (RF) algorithm, have emerged as promising tools for ET_c modeling due to their capacity to manage non-linear interactions among meteorological variables and enhance model flexibility. This study evaluates the performance of three ET_c estimation approaches for greenhouse-grown cucumber: the conventional FAO 56 K_c method, a non-linear K_c model using a third-degree polynomial, and direct ET_c prediction through the RF algorithm. These methods are assessed across two growth cycles, autumn-winter (A-W) and spring-summer (S-S), to capture seasonal differences in crop water requirements.

Materials and Methods

The study was conducted in a research greenhouse located at the College of Agriculture and Natural Resources, University of Tehran, focusing on daily ET_c of cucumber over two distinct growth periods. Environmental parameters were measured both inside and outside the greenhouse, including maximum, minimum, and average temperatures, relative humidity, and solar radiation. Reference evapotranspiration inside the greenhouse (ET_oG) was derived using a micro-lysimeter installed with a turfgrass surface, while daily ET_c was measured using a soil water balance method, where soil moisture content was monitored daily across three experimental plots to ensure precision. ET_c calculations were performed through three modeling approaches. In the first approach, the FAO 56 K_c model estimated ET_c by applying fixed crop coefficients and multiplying them by ET_oG . Although this method has been widely applied in field conditions, its applicability to greenhouses is limited due to fixed K_c assumptions. In the second approach, a non-linear K_c model was developed using third-degree polynomial regression on K_c values calculated as the ratio of ET_c to ET_oG , capturing growth-stage-specific variations. In the final approach, the RF model directly predicted ET_c based on a broad range of meteorological inputs. To optimize the RF model, hyperparameters were tuned using Python's GridSearchCV tool, and data were split into training (70%) and testing (30%) sets to validate model performance. After initial RF modeling, a feature selection process using Permutation Feature Importance (PFI) was applied to identify the most influential variables, refining the RF model to the top four parameters.

Results and Discussion

The results highlighted seasonal variability in cumulative ET_c , with the S-S period exhibiting nearly double the ET_c of the A-W period due to higher ambient temperature and increased solar radiation. These findings underscore the necessity of dynamic ET_c models that can accommodate seasonal and environmental variations. The FAO 56 K_c method produced a mean RMSE of 0.915 mm/day across both growth cycles, demonstrating limitations in fixed K_c approaches under greenhouse conditions. The non-linear K_c model, with an average RMSE of 0.64 mm/day,

provided improved accuracy by adjusting K_c values across different growth stages, especially during mid-growth when water demand peaks. This improvement aligns with the premise that non-linear models can better capture the ET_c variability within controlled environments. The RF algorithm demonstrated superior accuracy and flexibility, outperforming both K_c -based models with R^2 values of 0.96 for the A-W period and 0.94 for the S-S period in training datasets, and with respective RMSE values of 0.365 mm/day and 0.57 mm/day in testing datasets. These results illustrate the RF model's capacity to accurately model ET_c by capturing complex, non-linear interactions among variables such as maxTG (maximum temperature inside the greenhouse), meanRHG (average relative humidity inside the greenhouse), and RadiationG (solar radiation inside the greenhouse) during the A-W period, with RadiationG and EToG emerging as key variables during the S-S period. By emphasizing critical seasonal drivers of ET_c , the RF model offers a robust alternative that adjusts to environmental changes without relying on static K_c values. This adaptability supports RF's potential as a powerful tool for ET_c estimation, accurately reflecting seasonal influences on greenhouse crop water needs.

Conclusion

The findings from this study demonstrate that the RF algorithm, when applied to ET_c modeling in greenhouse conditions, provides a flexible, high-accuracy alternative to traditional K_c methods. Unlike the FAO 56 K_c and non-linear K_c models, which rely on predefined or growth-stage specific coefficients, the RF approach enables direct ET_c prediction using real-time meteorological data. By optimizing input variables through feature selection, RF efficiently reduced the model complexity, focusing on the top four influential parameters while retaining high predictive accuracy. This reduction not only streamlines data collection requirements but also enhances the model's applicability in practical greenhouse operations. The results indicate that RF's capacity to model complex relationships among variables makes it especially suited for greenhouse environments, where precision irrigation is crucial for sustainable water management. Ultimately, this research underscores the importance of integrating machine learning techniques in ET_c estimation, providing greenhouse operators with adaptive, resource-efficient tools for managing water use in controlled agricultural settings.

Keywords: Soil water balance, Python, Reference Evapotranspiration, Nonlinear crop coefficient, Lysimeter

Article Type: Research Article

Acknowledgement

We are grateful to all colleagues who helped us in the preparation of this research

Conflicts of interest

The authors of this article declared no conflict of interest regarding the authorship or publication of this article.

Data Availability Statement

We have no permission to release data and codes.

Authors' contribution

Morteza Khoshsimaie chenar: Writing original draft preparation, Resources, Software; **Arash Tafteh:** Manuscript editing; **Niazali Ebrahimipak:** Supervision

*Corresponding Author, E-mail: Khoshsima.mortaza@ut.ac.ir

Citation: Khoshsima, Morteza., Tafteh, Arash., & Ebrahimipak, N. (2025). Modeling Greenhouse Cucumber Evapotranspiration Using Machine Learning: A Random Forest Approach Versus Traditional and Non-linear Crop Coefficients. *Water and Soil Management and Modeling*, 5(2), 69-87.

doi: 10.22098/mmws.2024.16164.1516

Received: 10 November 2024, Received in revised form: 08 December 2024, Accepted: 11 December 2024, Published online: 22 June 2025.

Water and Soil Management and Modeling, Year 2025, Vol. 5, No. 2, pp. 69-87.

Publisher: University of Mohaghegh Ardabili

© Author(s)





مدل سازی تبخیر-تعرق خیار گلخانه‌ای با استفاده از یادگیری ماشین رویکرد جنگل تصادفی در مقابل ضرایب گیاهی سنتی و غیرخطی

مرتضی خوش سیمای چنار^{۱*}، آرش تافته^۲، نیازعلی ابراهیمی پاک^۳

^۱ دانشجوی دکتری آبیاری و زهکشی، گروه آبیاری و آبادانی، دانشکده کشاورزی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران
^۲ دانشیار مؤسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران
^۳ دانشیار مؤسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران

چکیده

تخمین مناسب تبخیر-تعرق گیاه (ET_c) برای برنامه‌ریزی آبیاری کارآمد بسیار حیاتی است. این مطالعه با هدف ارزیابی و مدل سازی ET_c و ضریب گیاهی (K_c) خیار گلخانه‌ای در دو دوره رشد پاییز-زمستان ۱۴۰۱ (A-W) و بهار-تابستان ۱۴۰۲ (S-S) در گلخانه تحقیقاتی پردیس کشاورزی و منابع طبیعی دانشگاه تهران انجام شد. تبخیر-تعرق مرجع داخل گلخانه (ET_{oG}) با استفاده از میکرو لایسمتر چمن و ET_c با استفاده از روش بیلان آب خاک و اندازه‌گیری روزانه رطوبت خاک در سه کرت آزمایشی انجام شد. K_c روزانه خیار گلخانه‌ای از نسبت ET_c به ET_{oG} محاسبه شد. ابتدا مدل FAO56 برای برآورد K_c (FAO56) و ET_c حاصل از حاصل ضرب آن در ET_{oG} ($ET_c = K_c \text{ FAO56} \times ET_{oG}$) مورد ارزیابی قرار گرفت که میانگین NRMSE و ME مربوط به ET_c در دو دوره رشد مورد بررسی به ترتیب برابر ۳۱/۲ درصد و ۰/۶۲ به دست آمد. در مرحله دوم از یک معادله درجه سه غیرخطی به منظور برآورد K_c و ET_c حاصل از حاصل ضرب آن در ET_{oG} ($ET_c = K_c \text{ NonLinear} \times ET_{oG}$) استفاده شد که میانگین NRMSE و ME در این روش نسبت به روش قبلی بهبود یافته و به ترتیب برابر ۲۱/۵۸ درصد و ۰/۸۲ به دست آمد. در مرحله سوم از الگوریتم جنگل تصادفی (RF) به منظور مدل سازی مستقیم ET_c خیار گلخانه‌ای با استفاده از ET_{oG} و پارامترهای هواشناسی اندازه‌گیری شده در داخل و بیرون گلخانه استفاده شد. نتایج ET_c حاصل از الگوریتم RF نشان داد که این الگوریتم با داشتن انعطاف پذیری و دقت بالا توانست پیش‌بینی‌های قابل قبول تری از ET_c ارائه دهد. میانگین NRMSE و ME مربوط به الگوریتم RF در دو دوره رشد به ترتیب برابر با ۶/۸ درصد و ۰/۹۵ در مرحله آموزش و ۱۱/۳ درصد و ۰/۸۶ در مرحله آزمایش به دست آمد. در مرحله سوم با بررسی اهمیت ویژگی‌ها به روش Permutation Feature Importance (PFI) و انتخاب تنها چهار پارامتر اندازه‌گیری شده در گلخانه به عنوان ورودی RF می‌توان بدون نیاز به K_c خطی یا غیرخطی، ET_c را با دقت بالایی برآورد کرد. این تحقیق اهمیت استفاده از روش‌های یادگیری ماشین در مدل سازی ET_c گلخانه‌ای را نشان داده و می‌تواند به بهبود مدیریت آب در گلخانه‌ها کمک کند.

واژه‌های کلیدی: بیلان آب خاک، پایتون، تبخیر-تعرق مرجع، ضریب گیاهی غیرخطی، لایسمتر

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: Khoshsima.mortaza@ut.ac.ir

استناد: خوش سیمای چنار، مرتضی، تافته، آرش، و ابراهیمی پاک، نیازعلی (۱۴۰۴). مدل سازی تبخیر-تعرق خیار گلخانه‌ای با استفاده از یادگیری ماشین: رویکرد جنگل تصادفی در مقابل ضرایب گیاهی سنتی و غیرخطی. *مدل سازی و مدیریت آب و خاک*، ۵(۲)، ۶۹-۸۷.

doi: 10.22098/mmws.2024.16164.1516

تاریخ دریافت: ۱۴۰۳/۰۸/۲۰، تاریخ بازنگری: ۱۴۰۳/۰۹/۱۸، تاریخ پذیرش: ۱۴۰۳/۰۹/۲۱، تاریخ انتشار: ۱۴۰۴/۰۴/۰۱

مدل سازی و مدیریت آب و خاک، سال ۱۴۰۴، دوره ۵، شماره ۲، صفحه ۶۹ تا ۸۷

© نویسندگان

ناشر: دانشگاه محقق اردبیلی



۱- مقدمه

صنعت گلخانه به دلیل بهره‌وری بالای آب و تولید محصول بسیاری از نقاط جهان گسترش یافته است. در دهه‌های اخیر، کشور ایران شاهد شتاب قابل توجهی در توسعه گلخانه‌ها بوده است. به‌ویژه پس از برنامه توسعه ده ساله که سطح زیر کشت گلخانه از ۶۰۲۱ هکتار در سال ۱۳۹۱ به ۱۴۱۴۳ هکتار در سال ۱۴۰۱ افزایش یافت (Jafari najafabadi et al., 2022). بر اساس آمارنامه‌های وزارت جهاد کشاورزی، محصولات گلخانه‌ای در سال‌های اخیر ۹/۸-۱۱/۹ درصد از کل تولیدات باغی کشور را تشکیل داده است، همچنین محصول خیار (*Cucumis sativus* L.) به‌عنوان یکی از مهم‌ترین محصولات گلخانه‌ای در کشور بوده و حدود ۵۷ درصد از سطح زیر کشت گلخانه‌ای را به خود اختصاص داده است. سبزیجات کشت شده در گلخانه‌ها به آب زیادی نیاز دارند و آبیاری تنها منبع آب برای رشد گیاه است و تبخیر-تعرق گیاه (ET_c) منبع اصلی مصرف آب آبیاری است. بنابراین، تخمین دقیق ET_c برای برنامه‌ریزی آبیاری کارآمد برای گیاهان گلخانه‌ای بسیار حیاتی است (Huang et al., 2020; Jafari najafabadi et al., 2022; Yan et al., 2024). در حال حاضر، مدل‌های متعددی برای تخمین ET_c در مزارع توسعه یافته و به کار گرفته شده‌اند (Allen et al., 1998; Priestley & Taylor, 1972). با این حال، به دلیل تفاوت‌های ساختاری و محیطی بین گلخانه‌ها و مزارع فضای باز، اکثر مدل‌های توسعه یافته ET_c برای کاربرد در گلخانه‌ها محدودیت دارند. مانند روش تشت تبخیر (K_p -method)، روش ضریب گیاهی (K_c -method)، مدل پنمن-مانتیث^۱ (r_c -method) و مدل پرستلی-تیلور^۲ (α_{PT} -method) (Priestley and Taylor, 1972; Allen et al., 1998; Gong et al., 2021; Nikolaou et al., 2023).

روش K_p به دلیل سادگی و سازگاری بالای آن با تغییرات آب و هوا، به مفیدترین روش برای تعیین ET_c تبدیل شده است (Tabari et al., 2013). تخمین قابل اعتماد ET_c با استفاده از روش K_p به تعیین دقیق ضریب تشت تبخیر (K_p) بستگی دارد که معمولاً با استفاده از روابط تجربی توسعه یافته مبتنی بر داده‌های حاصل از تشت تبخیر کلاس A تعیین می‌شود. به دلیل تفاوت‌های موجود در انواع تشت‌های تبخیر و شرایط کاربرد، معادلات متعددی برای تخمین K_p توسعه یافته‌اند. (Snyder, 1998; Pereira et al., 1995; Allen et al., 1998). تا به امروز، هیچ توافقی در مورد این که کدام روابط برای به‌کارگیری در گلخانه‌ها مناسب‌تر هستند، وجود نداشته است. کاربرد روش K_p نیازمند اندازه‌گیری مداوم تبخیر از تشت (E_{pan}) است. روش K_c

مبتنی بر عوامل هواشناسی، گیاهی و آبیاری، برای تعیین ET_c بر اساس تخمین تبخیر-تعرق مرجع (ET_0) زمانی که اندازه‌گیری‌های E_{pan} در دسترس نیست، عملی‌تر خواهد بود (Sedaghat, Ebrahimipak, et al., 2023; Sedaghat, Taftteh, et al., 2023). در روش K_c ، ضریب گیاهی به جزء ضریب گیاهی پایه (K_{cb}) و ضریب مربوط به تبخیر از سطح خاک (K_e) تقسیم می‌شود. K_{cb} به نوع گیاه، پارامترهای هواشناسی، رطوبت خاک و عملیات زراعی بستگی دارد (Allen et al., 1998). اگرچه Allen et al. (1998) مقادیر K_{cb} را برای گیاهان مختلف تحت اقلیم‌های مختلف ارائه کرده‌اند؛ اما بسیاری از مطالعات قبلی نشان دادند که مقادیر این ضریب برای یک گیاه ممکن است به دلیل تفاوت در شرایط آب و هوایی، بافت خاک و مدیریت‌های زراعی مختلف، به‌طور قابل توجهی متفاوت باشد (Ding et al., 2022; Jafari najafabadi et al., 2013). همچنین در شرایط استفاده از ضریب گیاهی یک جزئی، در نظر گرفتن ضریب ثابت در مراحل مختلف رشد، باعث کاهش دقت مدل در برآورد ET_c می‌شود.

یک راه جایگزین برای محاسبه مستقیم ET_c ، روش اصلی r_c (متفاوت از مدل FAO 56) با پارامتره کردن مقاومت آیرودینامیکی^۳ و مقاومت پوشش گیاهی^۴ است (Allen et al., 1998). مقاومت آیرودینامیکی را می‌توان با روش پیشنهادی Thom (1972) بر اساس سرعت باد، ارتفاع گیاه و توزیع پوشش گیاهی تعیین کرد. با این حال، ممکن است زمانی که سرعت باد نزدیک به صفر باشد (شرایط داخل گلخانه) مقاومت آیرودینامیکی بی‌نهایت شده و باعث خطا در برآورد ET_c شود (Gong et al., 2017). روش r_c نیازمند اندازه‌گیری دقیق پارامترهای هواشناسی و فرآیندهای پارامتره کردن است. برخلاف روش r_c ، روش آیرودینامیکی (α_{PT}) فقط به داده‌های دما و تابش نیاز دارد و به‌طور گسترده در مزارع فضای باز استفاده شده است (Priestley and Taylor, 1972; Gong et al., 2021). اما دقت روش α_{PT} عمدتاً به تعیین دقیق ضریب تجربی α بستگی داشته و بر اساس مطالعات انجام شده حتی در مراحل مختلف رشد سبزیجات گلخانه‌ای، مقادیر متفاوتی داشته و ثابت نیست (Gu et al., 2018; Gong et al., 2021).

مدل‌های ارائه شده برای پیش‌بینی ET_c محصولات گلخانه‌ای عمدتاً بر پایه روابط تجربی و استفاده از الگوریتم‌های ریاضی شکل گرفته‌اند. این روابط تجربی با بهره‌گیری از نظریه‌های مربوط به بیلان انرژی و پخشیدگی بخار آب به توسعه مدل‌های پیش‌بینی ET_c پرداخته‌اند، که عمدتاً تحت تأثیر عوامل هواشناسی

³ Aerodynamic resistance (r_a)⁴ Canopy resistance (r_c)¹ Penman-Monteith² Priestley-Taylor

و منابع طبیعی دانشگاه تهران، واقع در کرج انجام شد. سطح گلخانه ۸۰ مترمربع و پوشش سقف آن پلی کربنات با دیواره‌های شیشه‌ای و برای کشت نیز شمالی-جنوبی بود. آزمایش شامل سه کرت آزمایشی به ابعاد ۱ در ۸ متر و دو ردیف کاشت با فاصله ۵۰ سانتی‌متری و فاصله بوته‌های خیار در ردیف‌های کاشت ۴۰ سانتی‌متر بود. دوره رشدی پاییز-زمستان از ۲۷ مهر ۱۴۰۱ تا ۳۰ دی ۱۴۰۱ به مدت ۹۶ روز و بهار-تابستان از ۹ فروردین ۱۴۰۲ تا ۱۰ مرداد ۱۴۰۲ به مدت ۱۲۶ روز (از زمان انتقال نشاء به گلخانه) بود. طبق تعریف، لایسیمتر وزنی یا جعبه کشت یک تانک با ابعاد مشخص است که در داخل خاک قرار گرفته و لذا امکان اعمال رابطه بیلان جرمی آب که به صورت رابطه ۱ بیان می‌شود، در آن وجود دارد (James, 1988).

$$ET_c = I + P + SFI + LI + GW - RO - LO - L - DP - Z_r(\theta f - \theta i) \quad (1)$$

که در آن، ET_c تبخیر-تعرق گیاه کاشت شده در لایسیمتر بر حسب سانتی‌متر، I آبیاری بر حسب سانتی‌متر، P بارندگی بر حسب سانتی‌متر، SFI جریان سطحی ورودی به سطح خاک بر حسب سانتی‌متر، LI جریان زیرسطحی ورودی به حجم خاک بر حسب سانتی‌متر، GW مقدار آبی که از سفره‌های آب زمینی ممکن است وارد منطقه ریشه شود بر حسب سانتی‌متر، RO رواناب سطحی که از زمین خارج می‌شود بر حسب سانتی‌متر، LO جریان آب زیرسطحی که از زمین خارج می‌شود بر حسب سانتی‌متر، L نیاز آشوبی بر حسب سانتی‌متر، DP نفوذ عمقی بر حسب سانتی‌متر، Z_r عمق توسعه ریشه بر حسب سانتی‌متر، θf رطوبت حجمی خاک در شروع دوره موردنظر بر حسب سانتی‌متر مکعب بر سانتی‌متر مکعب و θi رطوبت حجمی خاک در انتهای دوره موردنظر بر حسب سانتی‌متر مکعب بر سانتی‌متر مکعب است. لایسیمتر از نظر هیدرولوژی بخش مجزا و کنترل‌شده‌ای از خاک است که پارامترهای LO ، LI و SFI در آن حذف شده و DP ، L ، RO و GW ، یا قابل اندازه‌گیری بوده یا حذف می‌شوند. هم‌چنین در گلخانه بارندگی وجود نداشته و می‌توان پارامتر P را نیز حذف کرد. بنابراین با اندازه‌گیری θf و θi می‌توان ET_c را با دقت بالا به دست آورد. استفاده از لایسیمتر برای محصولاتمانند خیار که نیاز به قیم دارند و وزن کل بوته روی نخ و قیم است دارای محدودیت بوده و بنابراین می‌توان از رابطه بیلان آب ساده شده، تبخیر-تعرق گیاه را اندازه‌گیری کرد. در این مطالعه از رابطه ۲ استفاده شد (James, 1988).

$$ET_c = I + P + Cr - Dp - Rf \pm \Delta S \quad (2)$$

که در آن ET_c تبخیر-تعرق گیاه بر حسب میلی‌متر، I مقدار آب آبیاری بر حسب میلی‌متر، P بارندگی بر حسب میلی‌متر، Cr

قرار دارند. این مدل‌ها به خاطر قابلیت‌های قوی و کاربرد گسترده، بسیار موردتوجه قرار گرفته‌اند؛ با این حال، این مدل‌ها دارای محدودیت‌های اشاره شده نیز هستند. به عنوان راه حل، با پیشرفت‌های تکنولوژیکی در زمینه کامپیوتر در سال‌های اخیر، بسیاری از پژوهشگران شروع به استفاده از روش‌های الگوریتمی ریاضی مدرن، مانند یادگیری ماشین، در مدل‌سازی ET_c کرده‌اند. الگوریتم جنگل تصادفی^۱ یکی از الگوریتم‌های یادگیری ماشین است که بر پایه ترکیب چندین درخت تصمیم^۲ عمل می‌کند و به‌ویژه در مسائل رگرسیون و دسته‌بندی بسیار کارآمد است. این الگوریتم با ایجاد نمونه‌های تصادفی از داده‌های آموزشی و ساخت درخت‌های تصمیم متعدد، قابلیت اطمینان و دقت پیش‌بینی خود را افزایش می‌دهد (Ignatenko et al., 2024). جنگل تصادفی به‌ویژه در مدل‌سازی ET_c کاربرد فراوانی دارد، زیرا می‌تواند به‌خوبی تعاملات غیرخطی و پیچیدگی‌های موجود در متغیرهای محیطی مانند دما، رطوبت و تابش را شناسایی کند (Chen et al., 2021; Ruiz-Alvarez et al., 2020). به‌منظور مدل‌سازی تبخیر-تعرق گوجه‌فرنگی گلخانه‌ای از الگوریتم جنگل تصادفی استفاده کرده‌اند. طبق نتایج ایشان، این الگوریتم با R^2 و RMSE به ترتیب ۰/۸۰۵ و ۰/۲۸۵ میلی‌متر در روز تبخیر-تعرق گیاه داخل گلخانه را با دقت خوبی برآورد کرده است. هم‌چنین طبق مطالعه Chen et al. (2020) الگوریتم جنگل تصادفی در مدل‌سازی و تفکیک تعرق گیاه کلم از تبخیر-تعرق گیاه در فضای باز دارای R^2 برابر ۰/۹۵ - ۰/۸۳ و RMSE ۰/۳۶ - ۰/۲۲ میلی‌متر در روز بوده که نشان دهنده دقت خوب و مناسب این الگوریتم است. به‌علاوه، این الگوریتم با دارا بودن قابلیت غربالگری ویژگی‌ها، می‌تواند متغیرهای مهم و تأثیرگذار بر ET_c را شناسایی کند؛ از این رو، به بهبود دقت پیش‌بینی‌ها کمک می‌کند. در نتیجه، جنگل تصادفی به عنوان ابزار قدرتمندی در تجزیه و تحلیل داده‌های اقلیمی و مدل‌سازی فرآیندهای هیدرولوژیکی گیاهان شناخته می‌شود (Li et al., 2020; Yong et al., 2023). با توجه به اهمیت برآورد دقیق ET_c در گلخانه و هم‌چنین اهمیت خیار گلخانه‌ای به عنوان بیش‌ترین محصول تحت کشت گلخانه‌های کشور، هدف از این مطالعه، واسنجی و بررسی روش‌های مختلف برآورد ET_c و K_c خیار گلخانه‌ای است.

۲- مواد و روش

۲-۱- آزمایش‌های گلخانه‌ای

این تحقیق به‌منظور تعیین و مدل‌سازی ET_c و K_c خیار گلخانه‌ای (رقم تالیسیا) در دو دوره رشدی پاییز-زمستان ۱۴۰۱ (A-W) و بهار-تابستان ۱۴۰۲ (S-S) در گلخانه تحقیقاتی پردیس کشاورزی

² Decision Tree

¹ Random Forest Algorithm (RF)

$$K_c = \frac{ET_c}{ET_0} \quad (5)$$

به منظور برازش خط مربوط به K_c بر اساس نشریه شماره ۵۶ فائو از رگرسیون قطعه‌ای^۱ در نرم افزار GraphPad Prism v. 10 طبق رابطه ۶ استفاده شده است.

$$\begin{cases} Y_{atDAT_0} = intercept_1 + slope_1 \times DAT_0 \\ Y_1 = intercept_1 + slope_1 \times DAT \\ Y_2 = Y_{atDAT_0} + slope_2 \times (-DAT_0 + DAT) \\ Y = if(DAT < X_0, Y_1, Y_2) \end{cases} \quad (6)$$

که در آن DAT روز پس از انتقال نشاء، DAT_0 تعداد روز پس از انتقال که گیاه به مرحله میانی رشد خود می‌رسد، Y_1 خط مربوط به K_c دوره توسعه گیاه، $intercept_1$ و $slope_1$ به ترتیب عرض از مبدأ و شیب خط مربوط به خط Y_1 و Y_2 ، K_c دوره میانی رشد، $slope_2$ شیب خط Y_2 که با توجه به ماهیت K_c ارائه شده در نشریه شماره ۵۶ فائو (Allen et al., 1998)، برابر صفر در نظر گرفته می‌شود. همچنین به منظور کاهش خطای مربوط به ثابت در نظر گرفتن K_c در طول دوره‌های رشد مختلف، K_c در طول دوره رشد بر اساس برازش یک رابطه درجه ۳ در زبان برنامه نویسی Python برآورد شد و در مدل سازی ET_c مورد استفاده قرار گرفت.

۲-۲- پارامترهای اندازه گیری شده

به منظور مدل سازی تبخیر-تعرق گیاه در گلخانه، دما و رطوبت نسبی هوا با استفاده سنسور دیجیتال مدل HTC-1 ساخت کشور چین (هر روز ساعت ۷/۳۰ صبح) اندازه گیری شد. حداقل، حداکثر و میانگین دما ($maxTG$ ، $minTG$ و $meanTG$) و رطوبت نسبی ($maxRHG$ ، $minRHG$ و $meanRHG$) و تابش داخل گلخانه ($RadiationG$) همانند تبخیر-تعرق مرجع (ET_0G) اندازه گیری شد و داده‌های هواشناسی بیرون گلخانه شامل حداقل، حداکثر و میانگین دما ($maxTO$ ، $minTO$ و $meanTO$) و رطوبت نسبی ($maxRHO$ ، $minRHO$ و $meanRHO$) و تبخیر-تعرق مرجع (ET_0O) در دو دوره رشد از سامانه نیاز آبی توسعه داده شده توسط موسسه تحقیقات خاک و آب کشور (داده‌های مربوط به ایستگاه هواشناسی کشاورزی واقع در ۳/۶ کیلومتری گلخانه و در مزرعه تحقیقاتی پردیس کشاورزی و منابع طبیعی دانشگاه تهران) استخراج شد. از داده‌های حاصل به منظور مدل سازی تبخیر-تعرق گیاه در گلخانه (ET_cG) استفاده شد. $RadiationG$ با استفاده از ضریب عبور نور (T) مربوط به جنس پوشش گلخانه ($T = 0.85$) از تابش بیرون گلخانه به دست آمد (Fernández et al., 2010). حداقل، حداکثر، میانگین و انحراف معیار پارامترهای استفاده شده در مدل سازی تبخیر-تعرق گیاه داخل گلخانه در شکل ۱ در دو دوره رشد A-W و S-S ارائه شده است.

صعود موئینه بر حسب میلی متر، Dp ، نفوذ عمقی بر حسب میلی متر، Rf ، رواناب بر حسب میلی متر و ΔS ، تغییرات رطوبت خاک در یک بازه زمانی مشخص در ناحیه ریشه گیاه است. در این مطالعه بارندگی (P) به دلیل شرایط گلخانه‌ای وجود ندارد. همچنین با توجه به سیستم آبیاری که در این مطالعه، آبیاری قطره‌ای است؛ از صعود موئینه، رواناب و نفوذ عمقی صرف نظر می‌شود. بنابراین رابطه بیان آب خاک به شکل رابطه ۳ درمی‌آید:

$$ET_c = I \pm \Delta S \quad (3)$$

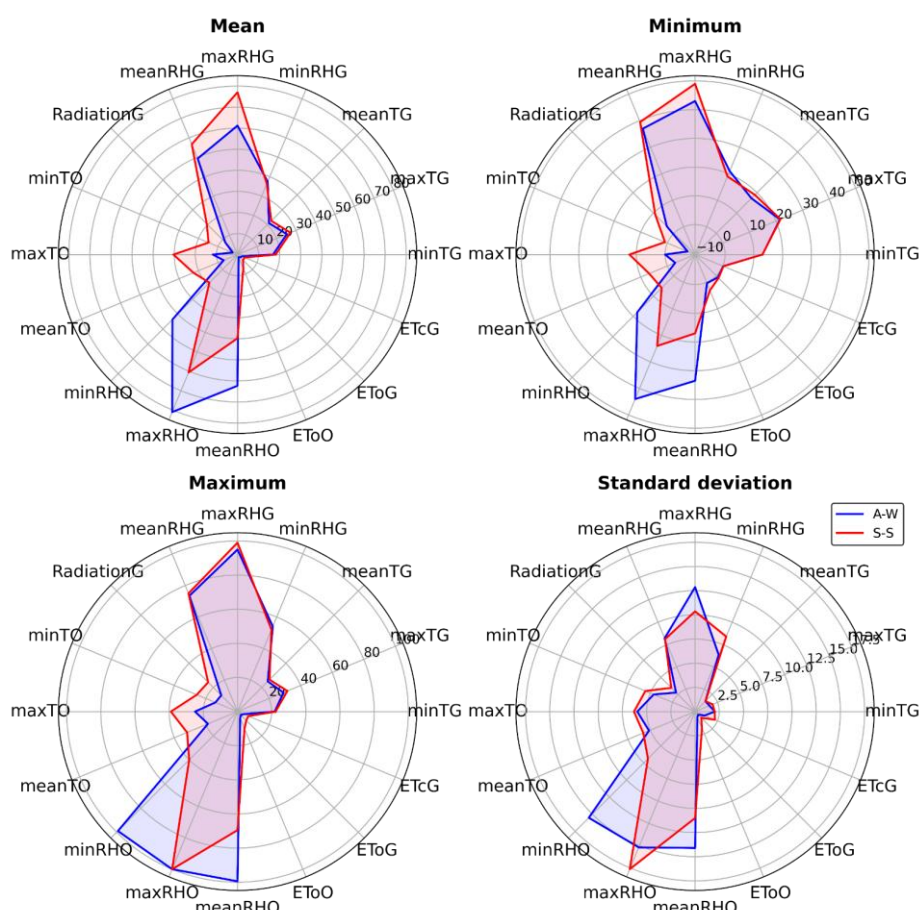
رطوبت خاک به صورت روزانه با استفاده از سنسورهای TDR دفن شده در لایه ۳۰-۰ سانتی متری اندازه گیری شد. از میانگین داده‌های رطوبت اندازه گیری شده در سه کرت آزمایشی بر اساس رابطه ۳ (رابطه بیان آب)، همچنین عمق ریشه در طول دوره رشد بر اساس رابطه ۲ برای تعیین تبخیر-تعرق گیاه خیار گلخانه‌ای (تغییرات رطوبت $\times$ در عمق ریشه) استفاده شد. عمق ریشه (Z_r) در طول دوره رشد با در نظر گرفتن حداکثر عمق ریشه ($Z_{r,m}$)، روزهای بعد از کاشت (DAP) و تعداد روزهایی که بعد از آن ریشه به عمق حداکثر می‌رسد (DTM)، از رابطه ۴ به دست آمد (Borg and Grimes, 1986; Merrill et al., 2002).

$$Z_r = Z_{r,m} \left(0.5 + 0.5 \times \sin \left[3.03 \left(\frac{DAP}{DTM} \right) - 1.47 \right] \right) \quad (4)$$

بر اساس منابع مطالعه شده، عمق ریشه خیار گلخانه‌ای تا یک متر هم می‌رسد اما عمق ریشه عموماً کمتر از ۳۵ سانتی متر است که در این مطالعه با توجه به عمق خاک بستر کاشت، $Z_{r,m}$ خیار ۳۰ سانتی متر در نظر گرفته شد. همچنین DTM ، ۳۵ روز در نظر گرفته شد (Rezvani et al., 2022).

در طول هر دو دوره رشد برای اندازه گیری روزانه تبخیر-تعرق مرجع (ET_0) از میکرو لایسیمتر چمن (ترازو با دقت ۲ گرم و ظرفیت ۴۰ کیلوگرم)، استفاده شد. کشت بذر چمن به طور یکنواخت و به مقدار ۷۰ گرم در هر متر مربع در میکرو لایسیمتر انجام گرفت. اندازه گیری‌های مربوط به محاسبه تبخیر-تعرق مرجع پس از آنکه ارتفاع چمن به ۱۲ سانتی متر رسید، آغاز شد. در طول دوره، ارتفاع چمن همواره بین ۱۱ تا ۱۳ سانتی متر ثابت نگهداری و برای جلوگیری از ایجاد تنش آبی همواره رطوبت داخل لایسیمتر در حالت ظرفیت زراعی و مناسب نگهداری شد. طبق تعریف نشریه شماره ۵۶ فائو، ضریب گیاهی (K_c) خیار گلخانه‌ای، در طول دوره رشد پاییز-زمستان (A-W) و بهار-تابستان (S-S)، طبق رابطه ۵ از نسبت تبخیر-تعرق گیاه (ET_c) به تبخیر-تعرق گیاه مرجع (ET_0) به دست آمد (Allen et al., 1998).

¹ Segmental regression



شکل ۱- حداقل، حداکثر، میانگین و انحراف معیار پارامترهای اندازه‌گیری شده در داخل و بیرون گلخانه به منظور مدل سازی تبخیر-تعرق خیار گلخانه‌ای در دو دوره رشد A-W و S-S

Figure 1- The minimum, maximum, mean and standard deviation of the parameters measured inside and outside the greenhouse in order to model the evapotranspiration of the greenhouse cucumber plant in the A-W and S-S growth periods

اما در الگوریتم RF که یک روش یادگیری گروهی مبتنی بر درخت‌های تصمیم است و هر درخت با استفاده از زیرمجموعه‌ای تصادفی از ویژگی‌ها ساخته می‌شود و پیش‌بینی نهایی، مجموعه‌های از پیش‌بینی‌های انجام شده توسط درختان است که بر مفروضات مشابه رگرسیون خطی تکیه نمی‌کند. برای جلوگیری از همبستگی درختان مختلف RF تنوع درختان را با رشد آن‌ها از زیرمجموعه داده‌های آموزشی مختلف که از طریق روشی به نام بسته‌بندی ایجاد شده‌اند؛ افزایش می‌دهد. بنابراین، پایداری بیش‌تری حاصل شده و در مواجهه با تغییرات جزئی در داده‌های ورودی، قوی‌تر شده و دقت پیش‌بینی را افزایش می‌دهد. به این ترتیب الگوریتم RF می‌تواند یک تخمین بی‌طرفانه از خطای تعمیم بدون استفاده از زیرمجموعه داده‌های خارجی محاسبه کند و به همین دلیل مشکل چندخطی در آن وجود ندارد و VIF در یک مدل جنگل تصادفی معنادار نیست (Breiman, 2001; Iraj, et al., 2024). بنابراین می‌توان نتیجه گرفت که هیچ‌گونه

۳-۲- الگوریتم جنگل تصادفی

روش‌های یادگیری ماشین ابزارهای قدرتمند پیش‌بینی تبخیر-تعرق در گلخانه و مزارع هستند. الگوریتم مورد استفاده در این مطالعه، الگوریتم جنگل تصادفی (RF) است. RF جزئی از روش‌های طبقه‌بندی و رگرسیون نظارت شده گروهی در حوزه داده‌کاوی است. در الگوریتم RF، داده‌های نمونه به تعداد n مجموعه تقسیم می‌شوند و از هر مجموعه، یک درخت برای آموزش استفاده می‌شود. درخت‌ها دارای یک عمق مشخص هستند و در یک گره، یک ویژگی به‌طور تصادفی انتخاب شده و بر اساس آن، تقسیم و شاخه‌بندی انجام می‌شود. با توجه به استفاده از چندین مجموعه نمونه، این الگوریتم مشکل داده‌های پرت را ندارد (Zhang et al., 2019). به منظور بررسی وجود چندخطی بین متغیرهای مستقل که در استفاده از مدل‌های رگرسیونی مشکل ساز بوده و باعث برازش بیش از حد می‌شود؛ از عامل تورم واریانس^۱ (VIF) استفاده می‌شود (Marcoulides and Raykov, 2018).

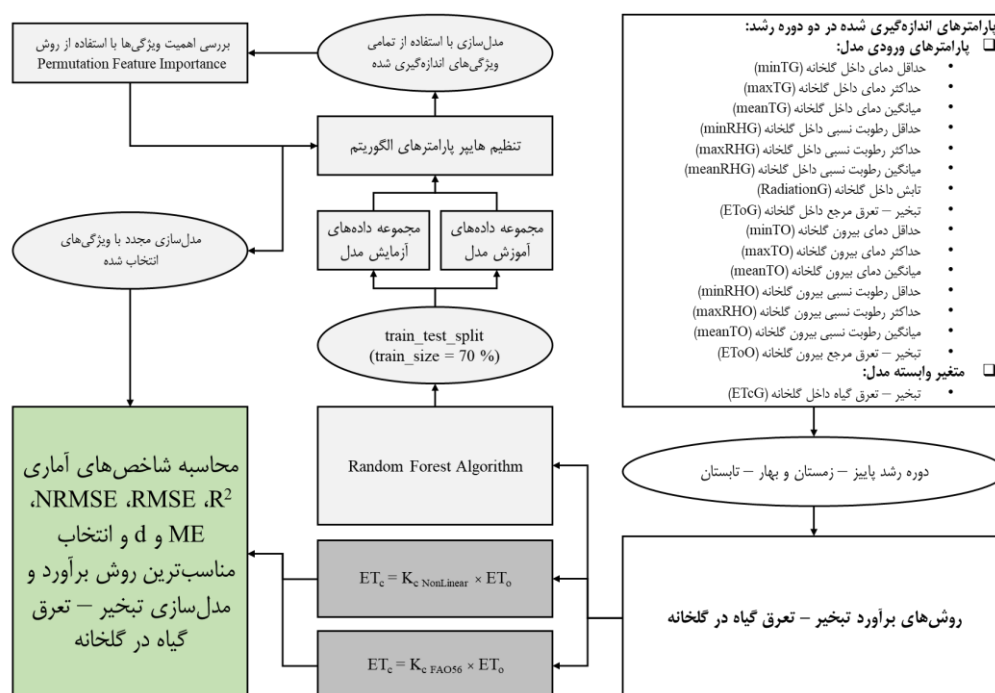
¹ Variance Inflation Factor, VIF

هم خطی و چندخطی بین متغیرهای مستقل انتخاب شده برای مدل سازی وجود ندارد. هاپیر پارامترهای^۱ الگوریتم RF، تعداد درخت^۲، حداکثر تعداد ویژگی ها^۳ (ورودی ها)، حداقل تعداد برگ های مورد نیاز برای تقسیم یک گره داخلی^۴، حداکثر گره های برگ در هر درخت^۵ و حداکثر عمق مجاز برای هر درخت تصمیم^۶ هستند که لازم است برای آن ها مقادیر مناسبی تعیین شوند (Breiman, 2001; Harris and Grunsky, 2015). در این مطالعه به منظور برآورد درست و مناسب هاپیر پارامترهای الگوریتم RF از تابع GridSearchCV در کتابخانه Scikit-learn استفاده شده است (Kramer, 2016).

۲-۴- مراحل مدل سازی

در این مطالعه، ET_c اندازه گیری شده در گلخانه به روش بیلان آب در دو دوره رشد A-W و S-S به عنوان پارامتر وابسته مدل و پارامترهای هواشناسی و تبخیر-تعرق مرجع اندازه گیری شده در داخل (ET_oG) و بیرون گلخانه (ET_oO) به عنوان پارامترهای ورودی مدل در نظر گرفته شدند. مدل سازی ET_c در دو دوره رشد به صورت جداگانه انجام گرفت. به منظور مدل سازی تبخیر-تعرق از سه روش K_c اندازه گیری شده بر اساس نشریه شماره ۵۶ فانو (K_c

برازش رابطه درجه ۳ بر روی داده های ضریب گیاهی اندازه گیری شده، K_c NonLinear) و الگوریتم RF (برآورد مستقیم ET_c بر اساس پارامترهای هواشناسی اندازه گیری شده در داخل و بیرون گلخانه) استفاده شد. برای توسعه مدل با استفاده از الگوریتم جنگل تصادفی، ابتدا تمام پارامترهای اندازه گیری شده در دو دوره رشد A-W و S-S به صورت جداگانه در قالب فایل اکسل آماده شده و سپس برای انجام مدل سازی از زبان برنامه نویسی Python استفاده شد. ۷۰ درصد تعداد نمونه های اندازه گیری شده به منظور آموزش^۷ مدل و ۳۰ درصد باقیمانده برای آزمایش و تست^۸ مدل در نظر گرفته شد (Kisi et al., 2006). برای تقسیم داده ها به مجموعه های آموزش و آزمایش از تابع train_test_split موجود در کتابخانه Scikit-learn استفاده شد (Kramer, 2016). هاپیر پارامترهای الگوریتم RF تنظیم شده و سپس با استفاده از تمام ویژگی های اندازه گیری شده در داخل و خارج گلخانه مدل سازی انجام شده و نتایج آن بررسی شد. سپس با استفاده از روش اهمیت ویژگی جایگشت^۹ (Kramer, 2016) ویژگی هایی (ورودی ها) که بیش ترین تأثیر را در مدل سازی داشته، مشخص شده و دوباره مدل سازی انجام شد. شکل ۲ مراحل مدل سازی در مطالعه حاضر را نشان می دهد.



شکل ۲- مراحل مدل سازی تبخیر-تعرق گیاه در گلخانه در دو دوره رشد A-W و S-S

Figure 2- Steps for modeling greenhouse crop evapotranspiration during the A-W and S-S growth periods

⁶ max_depth

⁷ Train

⁸ Test

⁹ Permutation Feature Importance, PFI

¹ Hyper parameters

² n_estimators

³ Max features

⁴ Min sample split

⁵ Max leaf nodes

۳-۱- تبخیر-تعرق گیاه اندازه‌گیری شده داخل گلخانه

تبخیر-تعرق روزانه و تجمعی اندازه‌گیری شده در گلخانه در دو دوره رشد A-W و S-S در شکل ۳ ارائه شده است. با توجه به نتایج، برای دوره رشد A-W حداقل، حداکثر، میانگین و مجموع تبخیر-تعرق گیاه خیار به ترتیب برابر ۰/۱۴ در ابتدای دوره رشد، ۳/۹۲، ۲/۷۲ میلی‌متر در روز و ۲۰۰ میلی‌متر به دست آمد. همچنین در دوره رشد S-S حداقل، حداکثر، میانگین و مجموع تبخیر-تعرق به ترتیب برابر ۰/۲۳ در ابتدای دوره رشد، ۶/۹، ۳/۷ میلی‌متر در روز و ۴۴۰/۱۱ میلی‌متر اندازه‌گیری شد. مشخص است که در دوره رشد S-S به دلیل طول دوره رشد بیش‌تر از A-W و همچنین افزایش تابش و دما، نیاز آبی گیاه افزایش پیدا کرده و نیاز آبی کل دوره رشد تقریباً دو برابر شده است. Rezvani et al. (2022) طول ۹۶ روز دوره رشد A-W، مجموع تبخیر-تعرق خیار گلخانه‌ای را ۲۱۴/۲ میلی‌متر گزارش کرده‌اند که با نتایج این تحقیق همخوانی دارد. اما در دوره رشد S-S، مجموع تبخیر-تعرق را ۲۲۲/۳ میلی‌متر برآورد کرده‌اند. این اختلاف را می‌توان به تفاوت در شرایط اقلیمی پیرامونی گلخانه، سازه گلخانه و استفاده یا عدم استفاده از تجهیزات کنترل اقلیم نسبت داد. همچنین Rezvani et al. (2022)، ET_c خیار را به روش Stanghellini (1987) به دست آورده‌اند. یکی از پارامترهای ورودی مدل Stanghellini (1987)، شاخص سطح برگ^۱ است که در مطالعه Rezvani et al. (2022) ایشان از یک رابطه رگرسیونی توسعه داده شده توسط Medrano et al. (2005) استفاده کرده‌اند و باعث این اختلاف شده است. Nikolaou et al. (2021) تبخیر-تعرق خیار گلخانه‌ای را برای دوره کشت هفت ماهه، ۵۶۶ میلی‌متر برآورد کردند. مقدار تبخیر-تعرق مرجع یا گیاه در گلخانه با توجه به رقم گیاه کشت شده، نوع سازه، نوع و سن پوشش که به میزان تشعشع ورودی به گلخانه مؤثر است، مدیریت و فصل کشت، شرایط اقلیمی منطقه و تجهیزات کنترل اقلیمی موجود در گلخانه، می‌تواند متفاوت باشد.

نتایج هر سه روش مدل سازی تبخیر-تعرق گیاه در گلخانه با استفاده از شاخص‌های آماری ضریب تبیین (R^2)، مجذور میانگین مربعات خطا (RMSE) و ریشه میانگین مربعات خطای نرمال شده (NRMSE)، کارایی مدل (ME) و شاخص توافق (d) به ترتیب طبق روابط ۷-۱۰ محاسبه شده (Moriasi et al., 2007) و پس از بررسی نتایج، مناسب‌ترین روش برآورد و پیش‌بینی تبخیر-تعرق گیاه در گلخانه انتخاب شد.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - S_i)^2}{n}} \quad (7)$$

$$NRMSE = RMSE \times \frac{100}{\bar{O}} \quad (8)$$

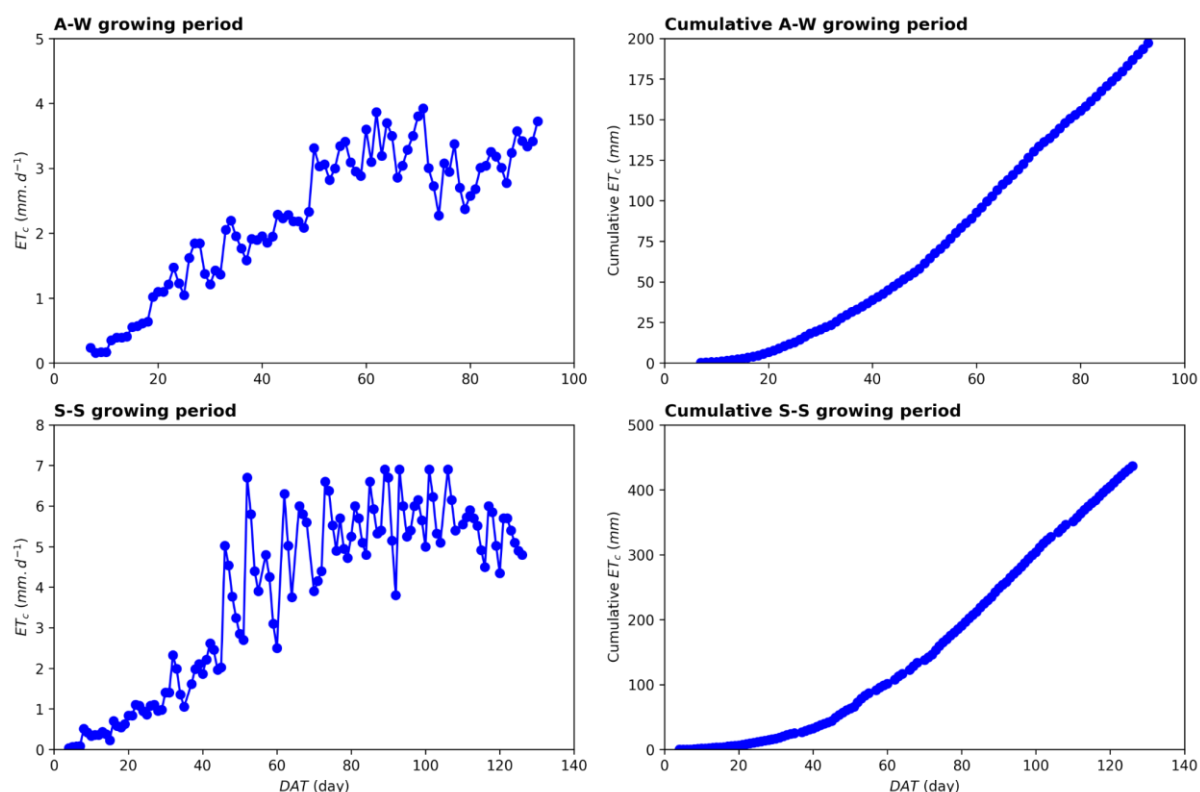
$$ME = 1 - \frac{\sum_{i=1}^n (O_i - S_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (9)$$

$$d = 1 - \frac{\sum_{i=1}^n (O_i - S_i)^2}{\sum_{i=1}^n (|S_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (10)$$

در رابطه‌های بالا O_i و S_i به ترتیب مقادیر پیش‌بینی مدل و اندازه‌گیری شده، n تعداد مقادیر مشاهداتی و $\bar{O}$ میانگین مقادیر اندازه‌گیری شده است. شاخص RMSE میزان تطابق منحنی مدل با داده‌ها یا فاصله پیش‌بینی‌ها با مقادیر اندازه‌گیری را اندازه می‌گیرد و هرچقدر نزدیک صفر باشد؛ مدل عملکرد خوبی داشته است. همچنین شاخص NRMSE نشان‌دهنده تفاوت نسبی بین داده‌های اندازه‌گیری و پیش‌بینی شده است. هنگامی که NRMSE کمتر از ۱۰، ۲۰-۱۰، ۳۰-۲۰ و بزرگ‌تر از ۳۰ درصد باشد، به ترتیب نشان‌دهنده عملکرد عالی، خوب، متوسط و ضعیف مدل در نظر گرفته می‌شود. شاخص ME نشان‌دهنده کارایی کلی مدل در مدل سازی بوده و می‌تواند از منفی بی‌نهایت تا ۱ باشد که عدد منفی و ۱ این شاخص به ترتیب نشان‌دهنده کارایی ضعیف مدل و عملکرد خوب آن است. شاخص d توافق بین داده‌های اندازه‌گیری و پیش‌بینی شده و اعداد نزدیک ۱ نشان‌دهنده عملکرد خوب مدل و توافق کامل داده‌های اندازه‌گیری و پیش‌بینی است (Jamieson et al., 1991; Taftah and Babazadeh, 2014).

۳- نتایج و بحث

¹ Leaf Area Index, LAI

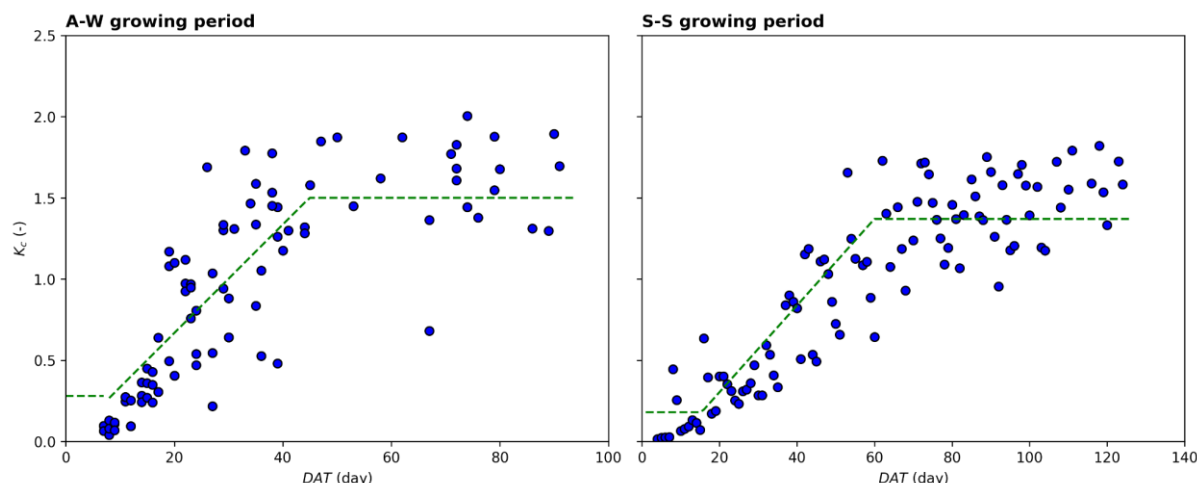


شکل ۳- تبخیر-تعرق روزانه و تجمعی خیار گلخانه‌ای در دو دوره رشد A-W و S-S
Figure 3- Daily and cumulative evapotranspiration of greenhouse cucumber during the A-W and S-S growth periods

A-W و S-S به ترتیب برابر ۰/۷۴ و ۰/۸۱ به دست آمده است. همچنین با توجه به ماهیت K_c ارائه شده در نشریه شماره ۵۶ فانو (Allen et al., 1998)، شیب خط Y_2 (slope₂) در رابطه ۶ برابر صفر در نظر گرفته شد. مطالعات مختلف نشان داده است که محصولات گلخانه‌ای که به صورت عمودی توسط نخ حمایت می‌شوند، عموماً دارای حداکثر مقادیر K_c بالاتری نسبت به محصولات مشابه در مزارع باز هستند (Orgaz et al., 2005; Incrocci et al., 2020). در مقابل، حداکثر مقادیر K_c برای محصولات گلخانه‌ای بدون حمایت عموماً مشابه مقادیر برای محصولات زراعی باز است. حداکثر مقدار K_c بیش‌تر محصولات سبزیجات گلخانه‌ای با حمایت عمودی، که معمولاً به ارتفاع ۱/۵-۲ متر می‌رسند، به سایبان‌های بازتر آن‌ها نسبت داده می‌شود که ناشی از بلندتر بودن و هرس منظم هستند. ترکیب سایبان‌های بازتر با تشعشعات در داخل گلخانه‌ها، جذب تشعشع و تعرق را افزایش می‌دهد (Orgaz et al., 2005).

۳-۲- مدل سازی تبخیر-تعرق بر اساس ضریب گیاهی

به منظور برازش خط مربوط به K_c از رگرسیون قطعه‌ای استفاده شده است که در شکل ۴ نتایج آن ارائه شده است. شکل کلی رگرسیون قطعه‌ای در رابطه ۶ ارائه شده است. در دوره رشد A-W طول دوره مرحله اولیه، توسعه و میانی رشد گیاه خیار گلخانه‌ای به ترتیب ۷، ۴۰ و ۴۷ روز بود. در این دوره با در نظر گرفتن DAT_0 برابر ۴۷ (اولین برداشت میوه)، با برازش خط توسط نرم افزار شیب خط Y_1 در مرحله توسعه برابر ۰/۳۳۴ و همچنین K_c در مرحله اولیه و میانی به ترتیب برابر ۰/۲۴ و ۱/۵ به دست آمده است. در دوره رشد S-S مرحله اولیه، توسعه و میانی رشد گیاه خیار گلخانه‌ای به ترتیب ۱۵، ۴۴ و ۶۷ روز بود و با در نظر گرفتن DAT_0 برابر ۶۰، با برازش خط توسط نرم افزار شیب خط Y_1 در مرحله توسعه گیاه برابر ۰/۲۶۶۹ و K_c در مرحله اولیه و میانی به ترتیب برابر ۰/۱۸ و ۱/۳۷ به دست آمده است. ضریب تبیین برای رگرسیون قطعه‌ای که در شکل ۴ ارائه شده است؛ در دو دوره رشد



شکل ۴- ضریب گیاهی خیار بر اساس نشریه FAO 56 در دو دوره رشد A-W و S-S

Figure 4- Crop coefficient of cucumber based on FAO 56 publication during the A-W and S-S growth periods

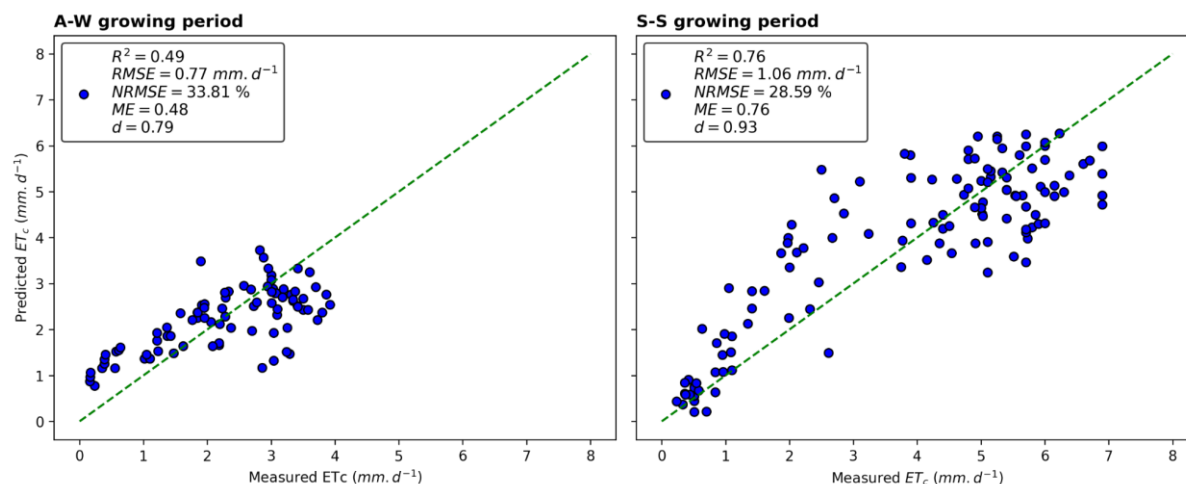
است. برای مرحله پایانی رشد خیار به دلیل جمع آوری کشت در خیار گلخانه‌ای به دلیل شرایط فصلی بازار و قیمت‌های پایین، قبل از این که مقدار K_c کاهش پیدا کند، عنوان کرده‌اند (Gallardo et al., 2013). در مطالعه Rezvani et al. (2022) در هر دو دوره رشد K_c میانی مشابه K_c به دست آمده در مطالعه حاضر است اما به صورت کلی، باید در نظر داشت که با توجه به زمان کشت، طول دوره رشد، تقسیم‌بندی مراحل رشد محصول، نوع سازه گلخانه، مدیریت کشت و اقلیم منطقه تفاوت‌هایی در این ضریب برای خیار گلخانه‌ای مشاهده می‌شود.

شکل ۵ مقایسه‌ای بین مقادیر اندازه‌گیری شده ET_c در داخل گلخانه و مقادیر پیش‌بینی بر اساس رابطه $ET_c = K_{c\text{FAO56}} \times ET_o$ برای دو دوره رشد A-W و S-S را نشان می‌دهد. در دوره رشد A-W و S-S شاخص‌های R^2 ، RMSE، NRMSE، ME و d به ترتیب برابر ۰/۴۹، ۰/۷۷، ۳۳/۸۱ درصد، ۰/۴۸ و ۰/۷۹ در دوره A-W و ۰/۷۶، ۰/۰۶، ۱/۰۶ میلی‌متر در روز، ۲۸/۵۹ درصد، ۰/۷۶ و ۰/۹۳ در دوره رشد S-S به دست آمد که نشان‌دهنده دقت پایین مدل در پیش‌بینی مقادیر تبخیر-تعرق در این دو دوره رشد مورد بررسی است. یکی از دلایل این عدم دقت می‌تواند ثابت و خطی فرض کردن K_c در دو مرحله ابتدایی و میانی رشد گیاه باشد. با مقایسه NRMSE، ME و d در دو دوره رشد، مشاهده می‌شود که مدل در دوره رشد S-S از دقت بالاتری در پیش‌بینی تبخیر-تعرق گیاه برخوردار است. شرایط محیطی نسبتاً پایدارتر و نور بیش‌تر در دوره S-S نسبت به دوره A-W ممکن است منجر به کاهش نوسانات و افزایش همبستگی بین مقادیر اندازه‌گیری و پیش‌بینی شده (R^2) شده باشد. در مقایسه با دوره A-W، احتمالاً ثابت و خطی در نظر گرفتن K_c تأثیر کم‌تری بر نتایج دارد، اما همچنان ممکن است منجر به خطاهایی در پیش‌بینی شود، به خصوص در مرحله میانی که گیاه در اوج فعالیت رشد و نیاز

مطالعات مختلفی به اندازه‌گیری K_c خیار گلخانه‌ای پرداخته‌اند. Blanco and Folegatti (2003) مقدار K_c خیار در یک گلخانه با پوشش پلاستیکی با ضخامت ۰/۱۵ میلی‌متر و در دوره رشد زمستان-بهار را در ۵، ۴۵ و ۱۰۵ روز پس از انتقال به ترتیب ۰/۱۹، ۱/۴۹ و ۱/۴۹ به دست آوردند که تأیید کننده نتایج این مطالعه است. Sun et al. (2019) ضریب گیاهی خیار گلخانه‌ای در کشت بهاره را در مرحله ابتدایی و میانی به ترتیب ۰/۴ و ۱/۱ گزارش کردند. Mushab (2020) در گلخانه‌ای تحقیقاتی در دانشگاه بغداد با پوشش پلی‌اتیلن ۲۰۰ میکرون، K_c خیار گلخانه‌ای را در مرحله آغازین، میانی و انتهایی فصل رشد به ترتیب ۰/۱، ۱/۲۹ و ۰/۹ گزارش کرد. همچنین طبق گزارش ایشان در یک مطالعه دیگر در دانشگاه بغداد، K_c در مرحله ابتدایی و میانی به ترتیب ۰/۱۶ و ۱/۲۳ به دست آمده است. Rezvani et al. (2022) در گلخانه‌ای در همدان، K_c خیار گلخانه‌ای (رقم ناگین) که بسیار شبیه رقم تالسیا است را در مرحله ابتدایی، میانی و انتهایی فصل رشد در کشت زمستانه به ترتیب ۰/۶۹، ۱/۴۳ و ۱/۰۵ و در کشت بهاره به ترتیب ۰/۶۳، ۱/۱۵ و ۰/۹ به دست آوردند. تفاوت در K_c اولیه مطالعه حاضر با مطالعه Rezvani et al. (2022) به دلیل تفاوت در نشاء می‌تواند باشد. در مطالعه حاضر، K_c نهایی یا مرحله پایانی (که معمولاً برای برخی از گیاهان در انتهای فصل رشد اندازه‌گیری می‌شود) محاسبه نشده است. این موضوع به دلیل تمرکز اصلی مطالعه بر روی مراحل رشد اولیه و میانی بوده است، که بیش‌ترین تأثیر را بر ET_c و در نتیجه بر نیاز آبی گیاه دارد. در بسیاری از موارد، K_c در مرحله پایانی به دلیل کاهش فعالیت رشد و آغاز رسیدگی میوه‌ها کم‌تر از سایر مراحل است و در شرایط گلخانه‌ای که دوره برداشت مستمر است، مرحله پایانی تعریف نمی‌شود. K_c خیار گلخانه‌ای در منطقه آلمریای اسپانیا در مرحله اولیه ۰/۲ و در مرحله میانی، ۱/۲ گزارش شده

گرفتن K_c در مراحل مختلف رشد، منجر به کاهش دقت پیش بینی ET_c می شود، زیرا K_c باید به تناسب شرایط رشد و مرحله فیزیولوژیک گیاه تغییر کند. بر اساس مطالعات، استفاده از K_c ثابت تنها در صورتی دقیق است که شرایط محیطی و نیاز آبی گیاه ثابت بمانند که در شرایط واقعی به ندرت اتفاق می افتد (Blanco and Folegatti, 2003).

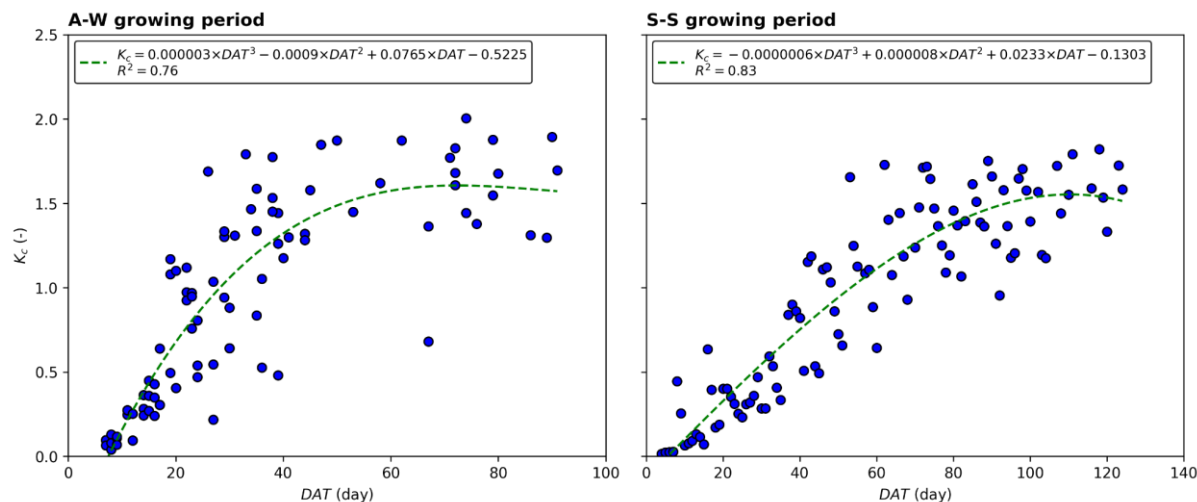
آبی است. همچنین، شرایط آب و هوایی متفاوت در دوره A-W مانند کاهش شدت نور و دما، می تواند بر میزان ET_c تأثیر بگذارد و تغییرات بیشتری را نسبت به دوره S-S ایجاد کند (Orgaz et al., 2005). این تغییرات دینامیک محیطی نشان می دهند که به جای استفاده از K_c ثابت، استفاده از ضریب های متغیر و وابسته به زمان برای دقت بیشتر توصیه می شود. ثابت و خطی در نظر



شکل ۵- تبخیر-تعرق اندازه گیری و پیش بینی شده خیار بر اساس ضریب گیاهی FAO 56 در دو دوره رشد A-W و S-S
Figure 5- Measured and predicted evapotranspiration of cucumber based on the FAO 56 crop coefficient during the A-W and S-S growth periods

درجه سه نمایش داده شده است. در هر دو دوره رشد، رابطه K_c برآزش داده شده به داده های اندازه گیری شده در شکل ۶ ارائه شده است. این مدل پیچیده تر نشان می دهد که K_c در مراحل مختلف رشد تغییر می کند و به اوج خود در مرحله میانی می رسد، سپس کاهش می یابد. مقدار R^2 برای دوره های رشد A-W و S-S به ترتیب برابر با ۰/۷۶ و ۰/۸۳ است که بیانگر تناسب خوب مدل با داده های اندازه گیری شده است.

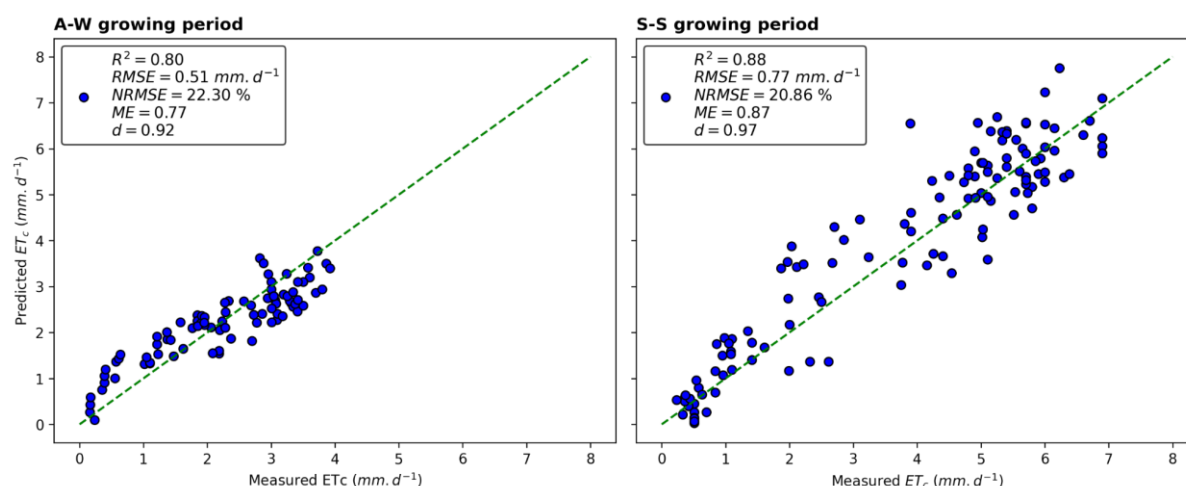
استفاده از یک K_c دینامیک یا مبتنی بر زمان، می تواند تأثیر به سزایی در بهبود دقت پیش بینی ها داشته باشد. تحقیقات نشان داده اند که تغییرات K_c بر اساس تغییرات فصلی و دوره های رشد می تواند نیاز آبی گیاه را بهتر منعکس کند و به این ترتیب خطاهای پیش بینی کاهش یابد (Sun et al., 2019). در شکل ۶ K_c خیار گلخانه ای به عنوان تابعی از تعداد روزهای پس از انتقال نشاء (DAT) به صورت غیرخطی (K_c NonLinear) و با استفاده از رابطه



شکل ۶- ضریب گیاهی غیرخطی خیار در دو دوره رشد A-W و S-S
Figure 6- Non-linear crop coefficient of cucumber during the A-W and S-S growth periods

و ۰/۹۷ در دوره رشد S-S به دست آمده است، که دقت بالای مدل K_c غیرخطی درجه سه نسبت به ضریب ثابت را تأیید می‌کند (شکل ۸). استفاده از این مدل غیرخطی به جای فرض خطی و ثابت برای K_c ، دقت بیشتری در برآورد ET_c ایجاد کرده است. مزیت اصلی استفاده از مدل درجه سه در این تحلیل، توانایی آن در به تصویر کشیدن تغییرات واقعی K_c در طول دوره رشد است. برخلاف مدل‌های خطی یا ثابت، این مدل می‌تواند تغییرات وابسته به مرحله رشد گیاه را در نظر بگیرد و از این رو دقت بیشتری در برآورد نیاز آبی گیاه دارد.

شکل ۷ مقایسه‌ای بین مقادیر اندازه‌گیری شده ET_c در داخل گلخانه و مقادیر پیش‌بینی بر اساس رابطه $ET_c = K_{c \text{ NonLinear}} \times ET_o$ برای دو دوره رشد A-W و S-S را نشان می‌دهد. استفاده از $K_{c \text{ NonLinear}}$ و برآورد تبخیر-تعرق با استفاده از آن توسط رابطه $ET_c = K_{c \text{ NonLinear}} \times ET_o$ باعث بهبود نتایج در برآورد ET_c در هر دو دوره رشد A-W و S-S شده است. شاخص‌های آماری محاسبه شده نیز نشان‌دهنده این موضوع است. R^2 ، RMSE، NRMSE، ME و d در دو دوره رشد A-W و S-S به ترتیب ۰/۸۰، ۰/۵۱ میلی‌متر در روز، ۲۲/۳۰ درصد، ۰/۷۷ و ۰/۹۲ در دوره رشد A-W و ۰/۸۸، ۰/۷۷ میلی‌متر در روز، ۲۰/۸۶ درصد، ۰/۸۷



شکل ۷- تبخیر-تعرق اندازه‌گیری و پیش‌بینی شده خیار بر اساس ضریب گیاهی غیرخطی در دو دوره رشد S-S و A-W
Figure 7- Measured and predicted evapotranspiration of cucumber based on non-linear crop coefficient during the autumn-winter (A-W) and spring-summer (S-S) growth periods

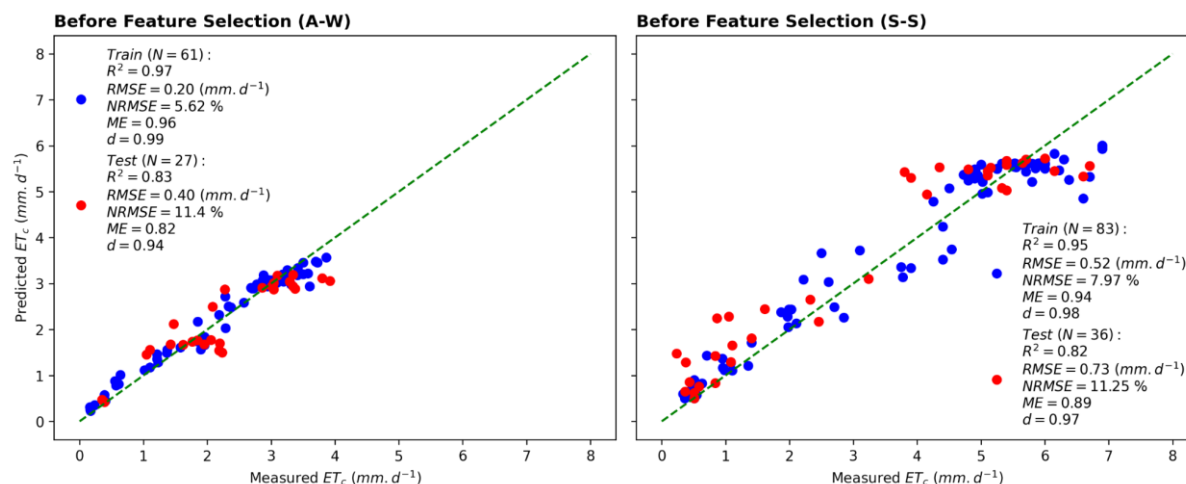
در روز، ۱۱/۴۰ درصد، ۰/۸۲ و ۰/۹۴ به دست آمده است. در دوره رشد S-S نیز مشاهده می‌شود که الگوریتم RF در هر دو مرحله آموزش و آزمایش عملکرد خوبی داشته و دارای R^2 ، RMSE، NRMSE، ME و d به ترتیب برابر ۰/۹۵، ۰/۵۲ میلی‌متر در روز، ۷/۹۷ درصد، ۰/۹۴ و ۰/۹۸ در مرحله آموزش و ۰/۸۲، ۰/۷۳، ۱۱/۲۵ درصد، ۰/۸۹ و ۰/۹۷ در مرحله آزمایش بوده است. الگوریتم RF به دلیل انعطاف‌پذیری بالا و قابلیت استفاده از داده‌های متنوع، یکی از روش‌های مؤثر در مدل‌سازی ET_c است. این الگوریتم از مجموعه‌ای از درخت‌های تصمیم استفاده می‌کند که هر کدام به طور مستقل از داده‌های مختلف و تصادفی تولید می‌شوند. این ویژگی باعث می‌شود که جنگل تصادفی بتواند روابط پیچیده و غیرخطی میان متغیرها را به خوبی شناسایی کرده و آن‌ها را در مدل‌سازی به کار گیرد (Breiman, 2001). یکی از مزایای این روش، انعطاف‌پذیری بالای آن در تطابق با داده‌های غیرخطی است که آن را به گزینه‌ای مناسب برای پیش‌بینی ET_c در مراحل مختلف رشد گیاه تبدیل می‌کند

۳-۳- تبخیر-تعرق گیاه بر اساس الگوریتم RF

شکل ۸ مقایسه‌ای بین مقادیر اندازه‌گیری شده ET_c در داخل گلخانه و مقادیر پیش‌بینی با استفاده از الگوریتم RF برای دو دوره رشد A-W و S-S را قبل از انتخاب ویژگی‌ها (کاهش تعداد ورودی‌های مدل) نشان می‌دهد. در استفاده از الگوریتم RF در هر دو دوره رشد، داده‌ها به دو بخش آموزش و آزمایش تقسیم‌بندی شده و شاخص‌های آماری ارائه شده در شکل ۸ در دو بخش آموزش و آزمایش الگوریتم است. بر اساس نتایج ارائه شده در شکل ۸ و برای دوره رشد A-W مشاهده می‌شود که الگوریتم RF با ۷۰ درصد داده‌ها به خوبی آموزش دیده است به طوری که شاخص R^2 ، RMSE، NRMSE، ME و d به ترتیب برابر با ۰/۹۷، ۰/۲، ۵/۶۲ درصد، ۰/۹۶ و ۰/۹۹ به دست آمده است و مقادیر اندازه‌گیری و پیش‌بینی شده ET_c تطابق بسیار خوبی دارند. این الگوریتم در مرحله آموزش نیز به خوبی عمل کرده و R^2 ، RMSE، NRMSE، ME و d به ترتیب برابر ۰/۸۳، ۰/۴ میلی‌متر

لحظه‌ای در شرایط محیطی به‌خوبی در مدل منعکس می‌شود و بدون نیاز به تنظیم دستی پارامترها، روابط پنهان میان متغیرهای هواشناسی و ET_c را کشف کرده و پیش‌بینی‌های دقیقی را ارائه دهد (Chen et al., 2020; Gong et al., 2020; Li et al., 2020).

(Ge et al., 2022). برخلاف روش‌های سنتی که از ضرایب گیاهی ثابت استفاده می‌کنند، الگوریتم جنگل تصادفی قادر است پارامترهای متنوع هواشناسی مانند دما، رطوبت نسبی و تابش خورشیدی را به‌عنوان ورودی مدل در نظر بگیرد. این امر دقت پیش‌بینی را به‌طور قابل‌توجهی افزایش می‌دهد، زیرا تغییرات



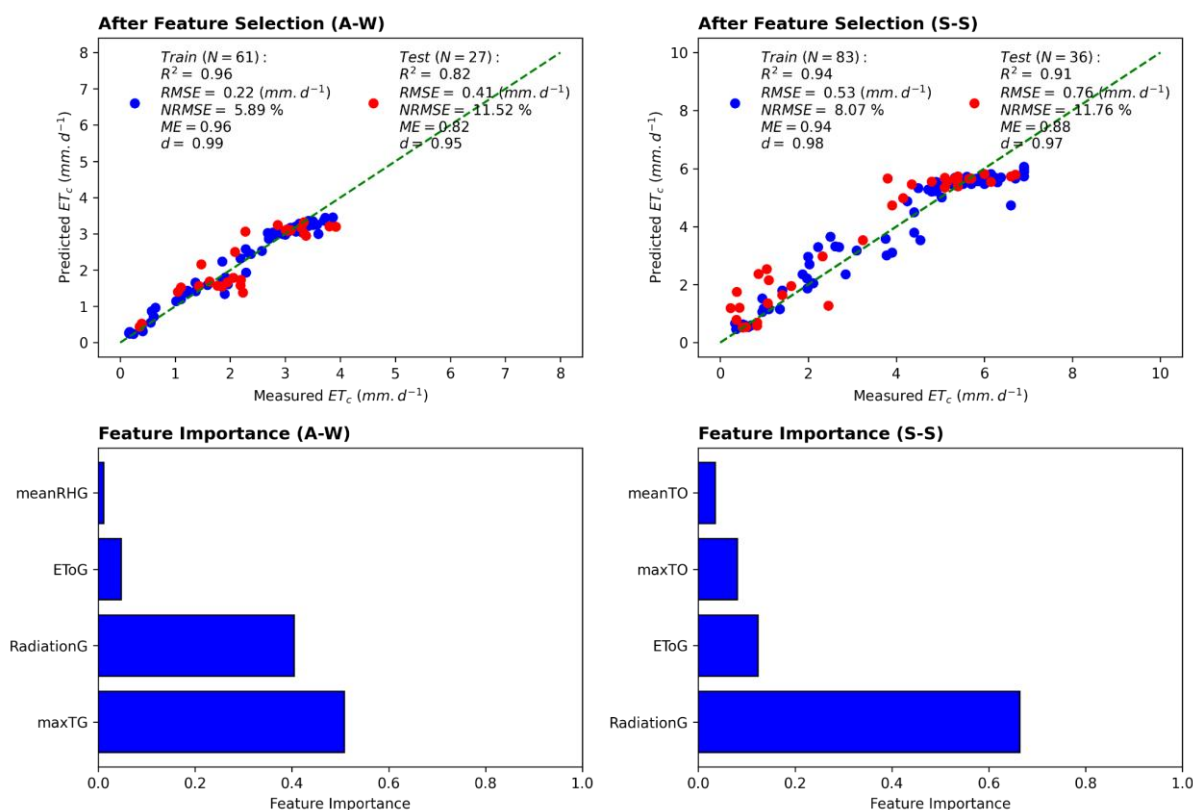
شکل ۸- تبخیر-تعرق اندازه‌گیری و پیش‌بینی‌شده خیار با الگوریتم RF در دو دوره رشد A-W و S-S و مرحله آموزش و آزمایش قبل از انتخاب ویژگی‌ها

Figure 8- Measured and predicted evapotranspiration of cucumber using the RF algorithm during the A-W and S-S growth periods and in training and testing phases before feature selection

یافت تا مدلی دقیق‌تر و ساده‌تر برای پیش‌بینی ET_c در گلخانه طراحی شود.

خروجی روش PFI در دو دوره رشد A-W و S-S در شکل ۹ ارائه شده است. با توجه به نتایج مشاهده می‌شود که پارامترهای ET_c ، $RadiationG$ ، $maxTG$ و $meanRHG$ در دوره رشد A-W و پارامترهای ET_c ، $RadiationG$ ، $maxTO$ و $meanTO$ در دوره رشد S-S به‌ترتیب بیش‌ترین تأثیر در ET_c را داشته و از اهمیت بالاتری برخوردار بودند. پارامتر $RadiationG$ در دوره رشد S-S (تابش داخل اندازه‌گیری شده در داخل گلخانه) در دوره رشد A-W از اهمیت بیش‌تری برخوردار بوده است اما در دوره رشد A-W رتبه دوم بعد از $maxTG$ (حداکثر دمای اندازه‌گیری شده در داخل گلخانه) قرار دارد. دلیل این موضوع را می‌توان به روشن بودن بخاری تابشی داخل گلخانه نسبت داد که در طول دوره سرما به‌دلیل کاهش تابش ورودی به گلخانه، سهم بیش‌تری در تأمین انرژی برای تبخیر-تعرق گیاه داشته است. همچنین در فصل سرما پارامتر $meanRHG$ (میانگین رطوبت نسبی اندازه‌گیری شده در داخل گلخانه) کنترل‌کننده میزان ET_c داخل گلخانه بوده است. اما در دوره رشد S-S که میزان تابش بالا بوده، این پارامتر از اهمیت بالاتری برخوردار بوده و انرژی موردنیاز برای تبخیر-تعرق گیاه را تأمین کرده است.

در مدل‌سازی ET_c با استفاده از الگوریتم‌های یادگیری ماشین، اغلب نیاز است پارامترهای متعددی به مدل اضافه شوند تا دقت پیش‌بینی افزایش یابد. این امر می‌تواند چالش‌های بیش‌تری را در فرآیند اندازه‌گیری‌های مشاهده‌ای ایجاد کند. بنابراین، کاهش تعداد پارامترها بدون افت دقت مدل، به یکی از چالش‌های اساسی در کاربرد مدل‌های یادگیری ماشین تبدیل می‌شود. این تکنیک با حذف برخی ورودی‌های مدل، متغیرهای اصلی و مؤثر بر دقت پیش‌بینی ET_c را شناسایی می‌کند تا مشخص شود آیا این حذف‌ها بر عملکرد مدل تأثیری دارند یا خیر. این فرآیند می‌تواند از پیچیدگی ناشی از وجود پارامترهای اضافه بکاهد. در این مطالعه، چندین پارامتر هواشناسی برای ET_c در نظر گرفته شد تا دقت مدل افزایش یابد. با این حال، هدف اصلی بررسی امکان کاهش پارامترهای ورودی با استفاده از روش PFI بود به‌طوری که همچنان مدلی کارآمد برای پیش‌بینی ET_c باشد. متغیرهای ورودی الگوریتم RF شامل حداقل، حداکثر و میانگین دما و رطوبت نسبی و تابش داخل گلخانه و تبخیر-تعرق مرجع اندازه‌گیری شده در داخل گلخانه و حداقل، حداکثر و میانگین دما و رطوبت نسبی و تبخیر-تعرق مرجع بیرون گلخانه بودند. در طی فرآیند بهینه‌سازی، تعداد پارامترهای ورودی به‌طور تدریجی کاهش



شکل ۹- تبخیر-تعرق اندازه‌گیری و پیش‌بینی شده خیار با الگوریتم RF و اهمیت ویژگی‌ها در دو دوره رشد A-W و S-S و مرحله آموزش و آزمایش بعد از انتخاب و کاهش ورودی‌ها

Figure 9- Measured and predicted evapotranspiration of cucumber using the RF algorithm and feature importance during the A-W and S-S growth periods and in training and testing phases after feature selection

جمع‌آوری داده‌ها را به شکل قابل‌توجهی کاهش داده و قابلیت کاربرد مدل در شرایط واقعی را افزایش داده است.

۴- نتیجه‌گیری

نتایج حاصل از این مطالعه نشان داد که استفاده از یک K_c ثابت در دوره‌های مختلف رشد منجر به افزایش خطای پیش‌بینی ET_c به‌ویژه در دوره‌های A-W می‌شود که تغییرات شرایط محیطی بیش‌تر است. برای بهبود دقت مدل‌های ET_c توصیه می‌شود از ضریب‌های گیاهی پویا و قابل‌تغییر بر اساس مرحله رشد و شرایط محیطی استفاده شود. استفاده از یک K_c دینامیک یا مبتنی بر زمان، می‌تواند تأثیر به‌سزایی در بهبود دقت پیش‌بینی ET_c داشته باشد. نتایج نشان داد که تغییرات K_c بر اساس تغییرات فصلی و دوره‌های رشد می‌تواند نیاز آبی گیاه را بهتر منعکس کند و به این ترتیب خطاهای پیش‌بینی کاهش یابد. الگوریتم جنگل تصادفی (RF) با بهره‌گیری از داده‌های هواشناسی و بدون نیاز به K_c ثابت یا حتی مدل‌های غیرخطی، می‌تواند با دقت بالایی ET_c خیار گلخانه‌ای را پیش‌بینی کند. کاهش تعداد پارامترهای ورودی از ۱۵ به چهار متغیر اصلی، علاوه بر بهینه‌سازی مدل، باعث کاهش

با توجه به نتایج ارائه شده در شکل ۹، مشاهده می‌شود که در هر دو دوره رشد A-W و S-S با کاهش تعداد پارامترهای ورودی و انتخاب این پارامترها برای مدل‌سازی ET_c خیار گلخانه‌ای، از دقت مدل کاسته نمی‌شود و همچنان از نظر شاخص‌های آماری (R^2 , $RMSE$, $NRMSE$, ME و d) که در شکل ۹ ارائه شده‌اند، برابر با مدل قبل از انتخاب ویژگی‌ها و روش PFI است. در این پژوهش، تعداد پارامترهای ورودی از ۱۵ متغیر به چهار متغیر اصلی کاهش یافت که این کاهش، گامی مؤثر برای ساده‌سازی و بهینه‌سازی مدل به شمار می‌آید. این کاهش با هدف حذف متغیرهای کم‌اهمیت و تمرکز بر پارامترهایی انجام شد که بالاترین تأثیر را بر دقت مدل پیش‌بینی تبخیر-تعرق داشتند. علاوه بر این، این رویکرد نشان دهنده قابلیت الگوریتم RF در برآورد مستقیم تبخیر-تعرق گیاه خیار در شرایط گلخانه‌ای است، به‌طوری‌که نیازی به استفاده از K_c خطی یا غیرخطی و نیز اجتناب از خطاها و دشواری‌های ناشی از اندازه‌گیری این ضرایب در برآورد ET_c نیست. در نتیجه، با حذف پارامترهای کم‌اهمیت، پیچیدگی مدل کاهش یافته و در عین حال، دقت پیش‌بینی آن حفظ شده است. همچنین، این کاهش تعداد پارامترها، زمان و منابع موردنیاز برای

ویرایش و بازبینی مقاله، کنترل نتایج؛ نیازعلی ابراهیمی‌پاک:
مشاوره، بازبینی متن مقاله

منابع

- ایرجی، مریم، موحدی نائینی، سید علیرضا، کمکی، چوقی بایرام، ابراهیمی، سهیلا و یغمایی، بامشاد (۱۴۰۳). ارزیابی پارامترهای مؤثر جهت پیش‌بینی عیار پتاسیم شورابه با استفاده از الگوریتم‌های ماشین بردار پشتیبان و جنگل تصادفی (مطالعه موردی: پلاهای شهرستان خور و بیابانک، استان اصفهان). *تحقیقات آب و خاک ایران*، (۱)۵۵، ۱۴۵-۱۶۱. doi: 10.22059/ijswr.2023.368909.669610
- جعفری نجف‌آبادی، محمد سعید، تافته، آرش و ابراهیمی پاک، نیازعلی (۱۴۰۱). تعیین نیاز آبی و آب کاربردی فلفل دلمه‌ای در گلخانه و مقایسه آن با نتایج سامانه نیاز آب. *تحقیقات آب و خاک ایران*، (۸)۵۳، ۱۸۳۱-۱۸۴۸. doi: 10.22059/ijswr.2022.345968.669321
- رضوانی، سید معین‌الدین، زارعی، قاسم و سالمی، حمیدرضا (۱۴۰۱). تبخیر-تعرق و ضریب گیاهی خیار گلخانه‌ای در منطقه همدان. *نشریه آبیاری و زهکشی ایران*، (۵)۱۶، ۹۰۴-۹۱۶. doi: 20.1001.1.20087942.1401.16.5.2.7
- صداقت، آزاده، ابراهیمی پاک، نیازعلی، تافته، آرش و حسینی، سید نرگس (۱۴۰۱). ارزیابی سه روش داده‌کاوی برای تخمین تبخیرتعرق مرجع در استان زنجان. *تحقیقات آب و خاک ایران*، (۱۲)۵۳، ۲۷۳۹-۲۷۵۷. doi: 10.22059/ijswr.2023.352890.669419
- صداقت، آزاده، تافته، آرش، ابراهیمی پاک، نیازعلی و حسینی، سید نرگس (۱۴۰۲). مقایسه برآوردهای تبخیرتعرق مرجع روزانه با روش‌های داده‌کاوی و سامانه نیاز آبی گیاهان در استان البرز. *هواشناسی کشاورزی*، (۲)۱۱، ۲۸-۱۶. doi: 10.22125/agmj.2023.336962.1134

References

- Allen, R. G., Pereira, L. S., Raes, D., Smith, M., & others. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome, 300*(9), D05109.
- Blanco, F. F., & Folegatti, M. V. (2003). Evapotranspiration and crop coefficient of cucumber in greenhouse. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 7(2), 285-291. doi: 10.1590/S141543662003000200017
- Borg, H., & Grimes, D. W. (1986). Depth Development of Roots with Time: An Empirical Description. *Transactions of the ASAE*, 29(1), 194-197. doi: 10.13031/2013.30125
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. doi: 10.1023/A:1010933404324/METRICS

پیچیدگی محاسبات و حفظ دقت پیش‌بینی شد. این مدل توانست هم در دوره رشد A-W و هم در S-S با شاخص‌های آماری بسیار قابل قبول از جمله مقادیر بالای R^2 و مقادیر کم RMSE و NRMSE در هر دو مرحله آموزش و آزمایش عملکرد موفقیت داشته باشد. کاربرد مدل RF به‌عنوان جایگزینی دقیق و منعطف برای روش‌های سنتی مبتنی بر K_c ، به‌ویژه در محیط‌های گلخانه‌ای با شرایط متغیر، می‌تواند به مدیریت بهینه آبیاری و کاهش مصرف آب کمک کند. نتایج این پژوهش، نقش روش‌های یادگیری ماشین در افزایش دقت و انعطاف‌پذیری پیش‌بینی‌های مرتبط با نیاز آبی گیاهان گلخانه‌ای را تأیید می‌کند و الگوریتم RF را به‌عنوان ابزاری مؤثر برای مدل‌سازی ET_c در این محیط‌ها و راهکاری مفید برای بهینه‌سازی مصرف آب در کشاورزی پیشنهاد می‌دهد.

سیاسگزاری

از تمامی همکارانی که در تهیه پژوهش حاضر ما را یاری نمودند کمال تشکر را داریم

تضاد منافع نویسندگان

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تضاد منفعی در خصوص نگارش و انتشار مطالب و نتایج این پژوهش ندارند.

دسترسی به داده‌ها

همه اطلاعات و نتایج در متن مقاله ارائه شده است.

مشارکت نویسندگان

- مرتضی خوش‌سیمای‌چنار: مفهوم سازی، انجام تحلیل‌های نرم‌افزاری/آماري، نگارش نسخه اولیه مقاله؛ آرش تافته: راهنمایی،
- Chen, H., Huang, J. J., & McBean, E. (2020). Partitioning of daily evapotranspiration using a modified shuttleworth-wallace model, random Forest and support vector regression, for a cabbage farmland. *Agricultural Water Management*, 228, 105923. doi: 10.1016/J.AGWAT.2019.105923
- Ding, R., Kang, S., Li, F., Zhang, Y., & Tong, L. (2013). Evapotranspiration measurement and estimation using modified Priestley-Taylor model in an irrigated maize field with mulching. *Agricultural and Forest Meteorology*, 168, 140-148. doi: 10.1016/J.AGRFORMET.2012.08.003
- Fernández, M. D., Bonachela, S., Orgaz, F., Thompson, R., López, J. C., Granados, M. R., Gallardo, M., & Fereres, E. (2010). Measurement and estimation of plastic greenhouse reference evapotranspiration in a Mediterranean climate. *Irrigation Science*, 28(6), 497-509.

- doi: 10.1007/S00271-010-0210-Z
- Gallardo, M., Thompson, R. B., & Fernández, M. D. (2013). Water requirements and irrigation management in Mediterranean greenhouses: the case of the southeast coast of Spain. *Good Agricultural Practices for Greenhouse Vegetable Crops; Plant Production and Protection Paper*, 217, 109–136.
- Ge, J., Zhao, L., Yu, Z., Liu, H., Zhang, L., Gong, X., & Sun, H. (2022). Prediction of Greenhouse Tomato Crop Evapotranspiration Using XGBoost Machine Learning Model. *Plants* 2022, Vol. 11, Page 1923, 11(15), 1923. doi: 10.3390/PLANTS11151923
- Gong, X., Liu, H., Sun, J., Gao, Y., Zhang, X., Jha, S. K., Zhang, H., Ma, X., & Wang, W. (2017). A proposed surface resistance model for the Penman-Monteith formula to estimate evapotranspiration in a solar greenhouse. *Journal of Arid Land*, 9(4), 530–546. doi: 10.1007/S40333-017-0020-8/METRICS
- Gong, X., Qiu, R., Ge, J., Bo, G., Ping, Y., Xin, Q., & Wang, S. (2021). Evapotranspiration partitioning of greenhouse grown tomato using a modified Priestley–Taylor model. *Agricultural Water Management*, 247, 106709. doi: 10.1016/J.AGWAT.2020.106709
- Gong, X., Wang, S., Xu, C., Zhang, H., & Ge, J. (2020). Evaluation of Several Reference Evapotranspiration Models and Determination of Crop Water Requirement for Tomato in a Solar Greenhouse. *HortScience*, 55(2), 244–250. doi: 10.21273/HORTSCI14514-19
- Gu, C., Ma, J., Zhu, G., Yang, H., Zhang, K., Wang, Y., & Gu, C. (2018). Partitioning evapotranspiration using an optimized satellite-based ET model across biomes. *Agricultural and Forest Meteorology*, 259, 355–363. doi: 10.1016/J.AGRFORMET.2018.05.023
- Harris, J. R., & Grunsky, E. C. (2015). Predictive lithological mapping of Canada's North using Random Forest classification applied to geophysical and geochemical data. *Computers & Geosciences*, 80, 9–25. doi: 10.1016/J.CAGEO.2015.03.013
- Huang, S., Yan, H., Zhang, C., Wang, G., Acquah, S. J., Yu, J., Li, L., Ma, J., & Opoku Darko, R. (2020). Modeling evapotranspiration for cucumber plants based on the Shuttleworth-Wallace model in a Venlo-type greenhouse. *Agricultural Water Management*, 228, 105861. doi: 10.1016/j.agwat.2019.105861
- Ignatenko, V., Surkov, A., & Koltcov, S. (2024). Random forests with parametric entropy-based information gains for classification and regression problems. *PeerJ Computer Science*, 10, e1775. doi: 10.7717/PEERJ-CS.1775
- Incrocci, L., Thompson, R. B., Fernandez-Fernandez, M. D., De Pascale, S., Pardossi, A., Stanghellini, C., Roupahel, Y., & Gallardo, M. (2020). Irrigation management of European greenhouse vegetable crops. *Agricultural Water Management*, 242, 106393. doi: 10.1016/j.agwat.2020.106393
- Iraji, M., Movahedi naeini, S. A., Komaki, C. B., Ebrahimi, S., & Yaghmaei, B. (2024). Evaluation of effective parameters for predicting the potassium grade of saline water by using support vector machine and random forest algorithms (case study: playa of Koor and Biabank area city, Isfahan province). *Iranian Journal of Soil and Water Research*, 55(1), 145–161. doi: 10.22059/ijswr.2023.368909.669610 [In Persian]
- Jafari najafabadi, M. S., Tafteh, A., & Ebrahimipak, N. (2022). Determining the Water Requirement and Applied Water of Bell Pepper in the Greenhouse and Comparing It with the Results of the Water Requirement System. *Iranian Journal of Soil and Water Research*, 53(8), 1831–1848. doi: 10.22059/ijswr.2022.345968.669321 [In Persian]
- James, L. (1988). *Principles of farm irrigation systems design*. <https://www.cabidigitallibrary.org/doi/full/10.5555/19891934086>
- Jamieson, P. D., Porter, J. R., & Wilson, D. R. (1991). A test of the computer simulation model ARCWHEAT1 on wheat crops grown in New Zealand. *Field Crops Research*, 27(4), 337–350. doi: 10.1016/0378-4290(91)90040-3
- Kisi, O., Karahan, M. E., & Şen, Z. (2006). River suspended sediment modelling using a fuzzy logic approach. *Hydrological Processes*, 20(20), 4351–4362. doi: 10.1002/hyp.6166
- Kramer, O. (2016). Scikit-Learn. *Studies in Big Data*, 20, 45–53. doi: 10.1007/978-3-319-33383-0_5
- Li, L., Chen, S., Yang, C., Meng, F., & Sigrimis, N. (2020). Prediction of plant transpiration from environmental parameters and relative leaf area index using the random forest regression algorithm. *Journal of Cleaner Production*, 261, 121136. doi: 10.1016/J.JCLEPRO.2020.121136
- Marcoulides, K. M., & Raykov, T. (2018). Evaluation of Variance Inflation Factors in Regression Models Using Latent Variable Modeling Methods. *Educational and Psychological Measurement*, 79(5), 874. doi: 10.1177/0013164418817803
- Medrano, E., Lorenzo, P., Sánchez-Guerrero, M. C., & Montero, J. I. (2005). Evaluation and modelling of greenhouse cucumber-crop transpiration under high and low radiation conditions. *Scientia Horticulturae*, 105(2), 163–175. doi: 10.1016/J.SCIENTA.2005.01.024
- Merrill, S. D., Tanaka, D. L., & Hanson, J. D. (2002). Root Length Growth of Eight Crop Species in Haplustoll Soils. *Soil Science Society of America Journal*, 66(3), 913–923. doi: 10.2136/SSSAJ2002.9130

- Moriasi, D. N., Arnold, J. G., Liew, M. W. Van, Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, 50(3), 885–900. doi: 10.13031/2013.23153
- Mushab, F. S. (2020). Forecasting Crop Coefficient Values for Cucumber Plant (*Cucumis sativus*). *Solid State Technology*, 63(6), 9085–9092. <https://www.solidstatetechnology.us/index.php/JSSST/article/view/5304>
- Nikolaou, G., Neocleous, D., Christou, A., Polycarpou, P., Kitta, E., & Katsoulas, N. (2021). Energy and Water Related Parameters in Tomato and Cucumber Greenhouse Crops in Semiarid Mediterranean Regions. A Review, Part I: Increasing Energy Efficiency. *Horticulturae* 2021, Vol. 7, Page 521, 7(12), 521. doi: 10.3390/HORTICULTURAE7120521
- Nikolaou, G., Neocleous, D., Kitta, E., & Katsoulas, N. (2023). Assessment of the Priestley-Taylor coefficient and a modified potential evapotranspiration model. *Smart Agricultural Technology*, 3, 100075. doi: 10.1016/J.ATECH.2022.100075
- Orgaz, F., Fernández, M. D., Bonachela, S., Gallardo, M., & Fereres, E. (2005). Evapotranspiration of horticultural crops in an unheated plastic greenhouse. *Agricultural Water Management*, 72(2), 81–96. doi: 10.1016/J.AGWAT.2004.09.010
- Pereira, A. R., Villa Nova, N. A., Pereira, A. S., & Barbieri, V. (1995). A model for the class A pan coefficient. *Agricultural and Forest Meteorology*, 76(2), 75–82. doi: 10.1016/0168-1923(94)02224-8
- Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 100(2), 81–92. doi: 10.1175/1520-0493(1972)100%3C0081:OTAOSH%3E2.3.CO;2
- Rezvani, S., Zarei, G., & Salemi, H. (2022). Evapotranspiration and crop coefficient of greenhouse cucumber in the Hamedan region. *Iranian Journal of Irrigation & Drainage*, 16(5), 904–916. https://idj.iaid.ir/article_159921.html [In Persian]
- Ruiz-Álvarez, M., Gomariz-Castillo, F., & Alonso-Sarria, F. (2021). Evapotranspiration Response to Climate Change in Semi-Arid Areas: Using Random Forest as Multi-Model Ensemble Method. *Water* 2021, Vol. 13, Page 222, 13(2), 222. doi: 10.3390/W13020222
- Sedaghat, A., Ebrahimipak, N., Taftah, A., & Hosseini, S. N. (2023). Evaluation of three data mining methods to estimate reference evapotranspiration in Zanjan province. *Iranian Journal of Soil and Water Research*, 53(12), 2739–2757. doi: 10.22059/ijswr.2023.352890.669419 [In Persian]
- Sedaghat, A., Taftah, A., Ebrahimipak, N., & Hosseini, S. N. (2023). Comparison of the reference evapotranspiration estimations by data mining methods and Crop Water Requirement System project in Alborz province. *Journal of Agricultural Meteorology*, 11(2), 16–28. doi: 10.22125/agmj.2023.336962.1134 [In Persian]
- Snyder, R. L. (1992). Equation for Evaporation Pan to Evapotranspiration Conversions. *Journal of Irrigation and Drainage Engineering*, 118(6), 977–980. doi: 10.1061/(ASCE)0733-9437(1992)118:6(977)
- Stanghellini, C. (1987). *Transpiration of greenhouse crops: an aid to climate management*. Wageningen University and Research.
- Sun, Y., Zhang, J., Wang, H., Wang, L., & Li, H. (2019). Identifying optimal water and nitrogen inputs for high efficiency and low environment impacts of a greenhouse summer cucumber with a model method. *Agricultural Water Management*, 212, 23–34. doi: 10.1016/J.AGWAT.2018.08.028
- Tabari, H., Grismer, M. E., & Trajkovic, S. (2013). Comparative analysis of 31 reference evapotranspiration methods under humid conditions. *Irrigation Science*, 31(2), 107–117. doi: 10.1007/S00271-011-0295-Z/TABLES/6
- Taftah, A., & Babazadeh, H. I. (2014). *Determine Yield Response Factors of Important Crops by Different Production Functions in Qazvin Plain*. <https://gnanaganga.inflibnet.ac.in:8443/jspui/handle/123456789/10474>
- Thom, A. S. (1972). Momentum, mass and heat exchange of vegetation. *Quarterly Journal of the Royal Meteorological Society*, 98(415), 124–134. doi: 10.1002/QJ.49709841510
- Yan, H., Zhao, S., Zhang, C., Zhang, J., Wang, G., Li, M., Deng, S., Liang, S., & Jiang, J. (2024). Calibration and assessment of evapotranspiration methods for cucumber plants in a Venlo-type greenhouse. *Irrigation and Drainage*, 73(1), 119–135.
- Yong, S. L. S., Ng, J. L., Huang, Y. F., & Ang, C. K. (2023). Estimation of Reference Crop Evapotranspiration with Three Different Machine Learning Models and Limited Meteorological Variables. *Agronomy* 2023, Vol. 13, Page 1048, 13(4), 1048. doi: 10.3390/AGRONOMY13041048
- Zhang, L., Huettmann, F., Zhang, X., Liu, S., Sun, P., Yu, Z., & Mi, C. (2019). The use of classification and regression algorithms using the random forests method with presence-only data to model species' distribution. *MethodsX*, 6, 2281–2292. doi: 10.1016/J.MEX.2019.09.035