

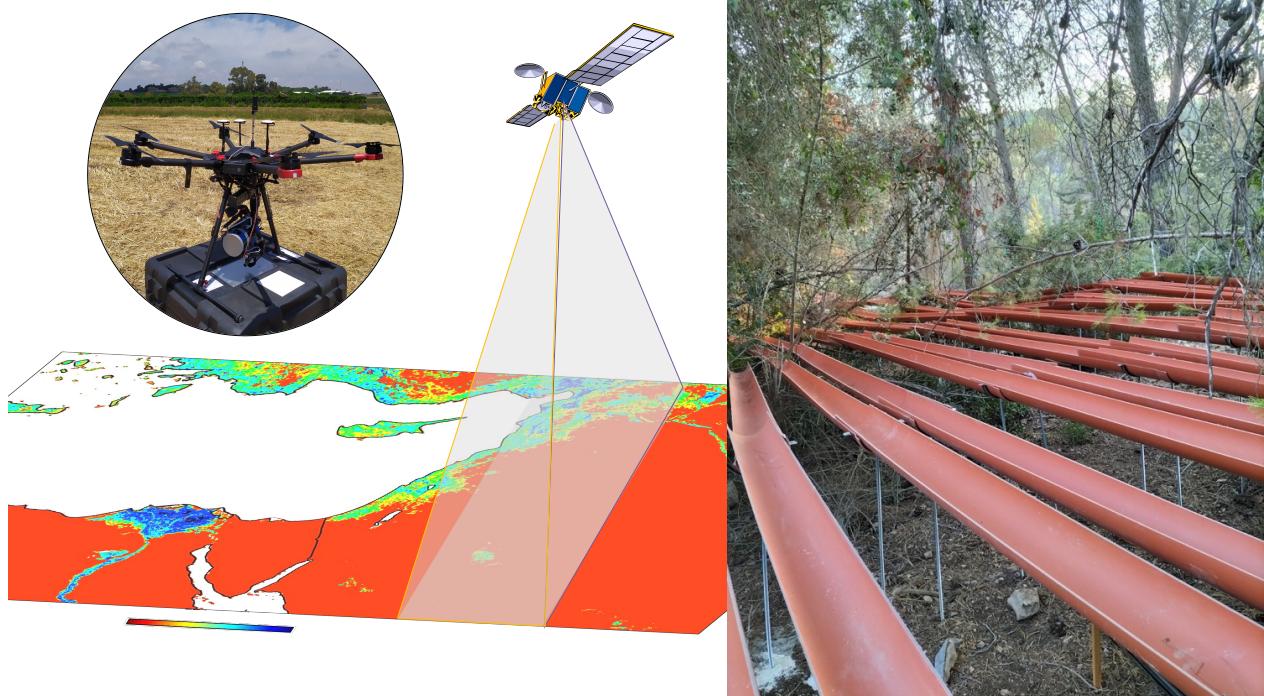


האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM



דו"ח מדעי מסכם

השפעת אקלים על פוטנציאל מים בעליים: בוחנת שיטות חישה מרוחק לעומת מדידות ישירות לאור שינוי אקלים



מחקר זה מומן על ידי:

הקרן הקיימת לישראל

פרויקט מס' 2145 (קק"ל 430/21)

25/06/25

כ"ט סיון התשפ"ה



האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM



דו"ח מדעי מסכם

השפעת אקלים על פוטנציאלי מים בעליים: בחינת שיטות חישה מרוחק לעומת מדידות שירות לאור שינוי אקלים

מוגש על ידי:

נתנאל פישמן¹, גבריאל מולרו^{1,2}, יהודה יונגשטיין^{1,2}, ירון מיכאל¹, גיל לרנר¹, אסף יעקובי³, סופי אוברסטינר³, לורה רץ³, תמייר קלין³ ודוד הלמן^{1,2}

¹מדעי הקרקע והמים, מכון למדעי הסביבה, הפקולטה לחקלאות, מזון וסביבה, האוניברסיטה העברית, רחובות

²בית הספר ללימודים סביבה מתקדמים, האוניברסיטה העברית בירושלים, קמפוס גבעת רם, ירושלים

³מחלקה למדעי הצמח והסביבה, מכון ויצמן, רחובות

מחקר זה מומן על ידי:

הקרן הקיימת לישראל

(פרויקט מס' 2145 (קק"ל 430/21)

25/06/25

כ"ט סיון התשפ"ה

תוכן עניינים

2	תקציר מנהליים
3	Executive Summary
4	התקדמות המחקר בשנה השלישי – 2024-2025
4	סיכום הממחקר
4	1. מבוא.....
4	2. פירוט מערכת הניסוי.....
6	2.1. תוצאות מדידות השדה.....
10	2.2. מודלי חיזוי פמ"ע - שיטות.....
15	2.3. תוצאות ניתוח מודלים.....
20	2.4. מודל סיווג.....
24	3. מסקנות, המלצות ממשקיות ופער ידע.....
25	מדד נתוניים
25	פירוט המשימות בהתאם לתוכנית המחקר
27	פרסומים
27	כנסים
27	נספח 1 – הפעלת אפליקציית מודל הסיווג
30	נספח 2 – מאמר 1 (Remote Sensing)
53	נספח 3 – מאמר 2 (Preprint)

תקציר מנהלים

שינויי האקלים באזורי הים-תיכוני יוצרים אתגרים הולכים וגוברים לעורות טבעיות ומיוירעים, בעיקר על רקע עליה בתדרות ובעוצמת אירועי בצורת. מחקר זה בוחן את האפשרות לנטר עקט יובש בעצים באמצעות שילוב בין מדידות שירות (פוטנציאלי מים בעליים – פמ"ע) לבין חישה מרוחק מרחפן ולויין. במסגרת המחקר, הוקם ניסוי הפחתת גשם בעיר ישע' שבפלת יהודה, תוך שימוש ברחפן היפרסתטראלי ובמערכות מדידה ייחודית לממד פיזיולוגית.

במהלך שלוש שנים של מחקר נאספו נתונים מגוון מקורות: מדידות פמ"ע עבור חmissה מימי עצים, דימותים היפרסתטראליים, נתונים תחנות מטאורולוגיות, וסימולציה של ערוצי לויין (Sentinel-2, Planet, S_NE_V). בשנה השלישית בוצעה סינתזה של המידע שנאסף, פותחו מודלים מבוססי למידת מכונה לזיתו מוקדם של עקט יובש, והורחבה הchèלה של המודלים לאתרים נוספים בישראל.

המצאים מראים כי:

- הפחתת המשקעים אכן הפחיתה משמעותית את תכולת המים בקרקע, אך לא גרמה לשינויים מובהקים במדדים הפיזיולוגיים (פמ"ע, דיות, הטמפרטут פחמן דו-חמצני) ברוב המינים.
- חישה היפרסתטראלית הצלחה לזהות חתימות ספקטראליות המעידות על עקה לפני הופעת סימנים פיזיולוגיים מובהקים.
- מודל לוגיסטי שהtabesse על נתונים רחפן הצליח לסוג עצים במצב עקה בדיק של 85% (Recall: 94%, F1: 83%)
- מודלים שאומנו על ערוצי הלויין ה-S_NE_V הצליחו לזהות מוקדי עקה באתרים נוספים, השנה עד שנתיים טרם נצפתה בפועל תמותת עומדים.

המסקנות המשמעות המרכזיות הן:

- חישה מרוחק היפרסתטראלית, בשילוב למידת מכונה, מהווה כלי חזק לזיתו מוקדם של עצים בסיכון תמותה.
- יש להטמייע יכולות אלו במסגרת ניטור עיר, תוך קידום פתרונות מבוססים על נתונים לוייניים.
- פותחה אפליקציה לזיתו אזרחי פגיעה עומדים בעקבות עקט-מים על בסיס דימותי לויין ומודל מבוסס למידת מכונה שפותח במסגרת מחקר זה. האפליקציה מיועדת לשימוש קק"ל לצורך זיתו אזרחי סיכון פוטנציאליים (הסביר וקיים לאפליקציה בנספח 1).
- מומלץ להמשיך את המחקר לאזורים נוספים, ולשלבו בתכנון ניפוי של טיפולים מניעי תמותת עיר.

Executive Summary

Climate change imposes growing stress on Mediterranean forests, particularly due to increasing drought frequency and severity. This study investigated the feasibility of detecting drought stress in trees using a combined approach that incorporates direct measurements (leaf water potential – LWP) and remote sensing from drones and satellites. The research included a rainfall exclusion experiment in Yishi Forest (Judean foothills), hyperspectral UAV imaging, and multi-source physiological and environmental monitoring.

Over the course of three years, the research team collected extensive field data, including LWP and LAI measurements for five dominant tree species, UAV hyperspectral imagery, meteorological station records, and simulated satellite reflectance (VENμS, Sentinel-2, Planet). In the third year, we synthesized all data, developed machine learning models for drought classification, and applied the methods to external forest sites.

Key findings:

- Rainfall exclusion reduced soil moisture by ~47%, but had limited physiological impact (LWP, assimilation, transpiration) in most species.
- Hyperspectral imaging detected pre-visual drought stress via spectral signatures not captured by physiological measurements.
- A logistic regression model based on UAV data achieved 85% accuracy (Recall: 94%, F1: 83%).
- Models trained on simulated VENμS bands successfully identified drought risk areas ~1 year prior to visible canopy decline at external forest sites.

Management-oriented conclusions:

- Hyperspectral remote sensing with ML is a powerful tool for early drought detection in forests.
- Integration of these technologies into forest monitoring systems is recommended.
- We developed a web application that provides early warnings on stands at risk following drought events. The app is based on machine learning models and spectral data derived from satellite imagery (VENμS or Sentinel-2) uploaded by the user (the link and explanations are provided in **Appendix 1**).
- Further upscaling and spatial expansion of this approach are needed for forest climate resilience planning.

התקדיםות המחקר בשנה השלישייה (והאחרונה) – 2024-2025

השנה השלישית למחקר הtmpקדה בניתוח הנתונים שנאספו, בניית מודלים, התאמתם לנוטוי לויין ופיתוח כלים ישומיים, לאחר שהמודדים הגיעו לשטח הסטיימנו בשנת 2023. התוצר המרכזי של המחקר היא אפליקציה המאפשרת סיוג בעזרת נתונים לויין של אזורים בעלי סיכון גבוה ונמוך להתיישבות.

השנה השלישית למחקר הchallenge מאוחר בעקבות עיכובים שהלו בשל פריצת מלחמת הרבות ברצל באוקטובר 2023, ויצאה למילואים של סטודנטים המעורבים במחקר.

סיכום המחקר

1. מבוא

שינויי אקלים עלולים להשפיע על הצומח הן מבחינת תפוקדו, והן מבחינת הקשר שבין מקורות המים. על אף הצלחה הכלכלית של מפעל הייצור בישראל, בשלושת העשורים האחרונים סבלו הערים בארץ (ובעולם) מאירועי תמותה נרחבים, אשר התרחשו על רקע התchmodות גלובלית ושנות בצורה רצפות. אף טרם התמונה, מראים עצים הסובלים מתקת-יובש ירידת תפוקודית, הפגעת בשירותי המערכת של העיר. למרות יכולתם של העצים להסתגל לשינויי אקלים, לרבות יובש, החשש הוא כי קצב השינויים יעלה על קצב ההסתגלות.

אחד המדרדים השימושיים ביותר לאפקון עקת-יובש עצים הינו פוטנציאל המים בעליים (פמ"ע), מפני שהוא מייצג את מידת הייבות של המים בציוריות ההובלה של העץ – כלומר, ערך זה נעשה שלילי יותר ככל שగברת עקת-היובש. לעומת זאת, ניתן למדוד פמ"ע בשדה, מדידה זו מוגבלת בכיסוי המרחב שלה, ובכך יכולה שלא להעיר השפעות מרחביות מעבר לקנה מידה מקומי. שימוש בכליל חישה מרוחק (חישנים ומצלמות על רחפנים ולויינים) יכול לספק מידע שבעזרתו ניתן יהיה להעיר את פמ"ע בקנה מידה מרחבית רציף וגדול יותר מאשר מחינת פריסת טחח. מידול פמ"ע, באמצעות נתוני השטח ונתוני החישה מרוחק, מאפשר לבחון שינויים בתנאי עקה במספר יעורים לאורך מפל הגשם בישראל, ולאורך מספר שנים אחרות.

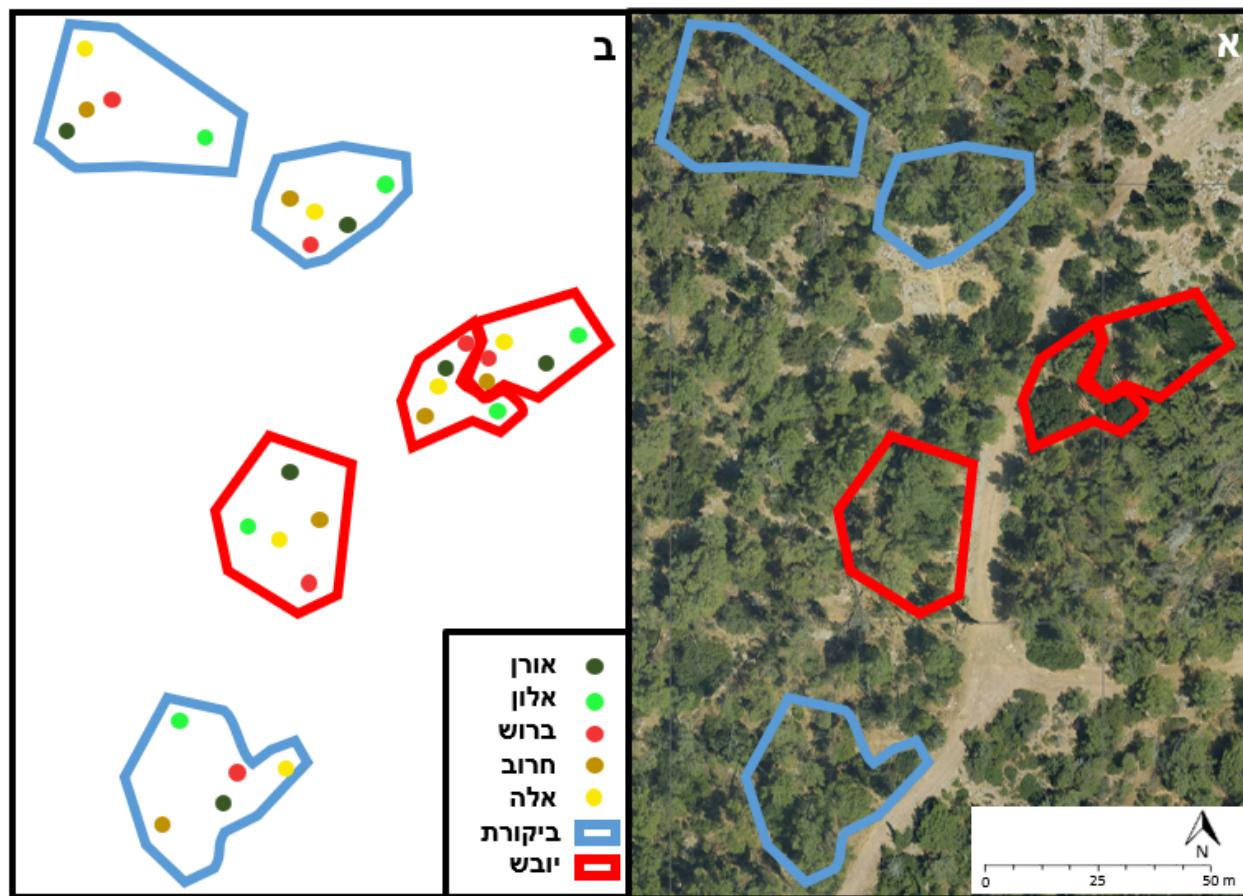
בכדי לאלץ תנאי עקת-יובש בחלוקת הניטור, הקמנו מערכת ניסוי של הפחתת גשם (הפחתה של כ-50%) אשר ידמה בצורת מתמשכת בעיר ישע. בחלוקת הביקורת וההפחתה, נמדד פמ"ע במגוון עצים באופן רציף (כל שבועיים-שלושה), יחד עם מדרדים נוספים המבטים את תפוקוד העצים כגון קצב קיבוע פד"ח (A) וקצב הדימות (Tz). זאת בנוסף לצלום החלקות בעזרת מצלמה היפרסקטרלית וממשק לדר, המותקנים על רחפן ה- So 600 Pro DJI שלו.

כל הנתונים שנאספו נוחחו הן במתודות מתמטיות פשוטות (ניתוחים סטטיסטיים ואינדקס NDSI) והן באמצעות שימוש באלגוריתמים של למידת מכונה, לייצור 2 סוגים מודלים: הראשון הוא מודל חיזוי פמ"ע, המאפשר מיפוי של פמ"ע בחופות עצי העיר באמצעות חישה מרוחק (בשלב הראשון ברמת הרחפן), ומtower ניסיון להסיק מסקנות גם לגבי רמת הלויין. המודל השני הוא מודל סיוג, מודל למידת מכונה הממיין את הפיקסלים של חותמת עצי העיר הנקלטים על ידי החישון לכאלו שנמצאים בעקת יובש וכאלו שאיןם, לייצור יצרת מפת סיכונים המהווה תומך-החלטה לניהול ממשק העיר.

2. פירוט מערכת הניסוי

כאמור, שלב הראשון של המחקר הzbוצע בעיר ישע, עיר ים תיכוני-מעורב בשפלת יהודה. המחקר בישע הtmpקدة ב-5 מיני מעוזים – אורן ירושלים, אלון מצוי, ברוש, אלת המסתיק וחרוב. המחקר הzbוצע במתכונת של הפחתת

גשם – כאשר ב-3 חלקיות מותקנת מערכת הפחתת גשם (הפחתה של כ-50%. השטח המסומן באדום, איור 1), ו-3 חלקיות המשמשות כביקורת (השטח המסומן בכחול, איור 1). המערכת כוללת חצאי צינור המותקנים בגובה של כ-1 מטר מעל הקרקע, אך מתחת לחופת העצים על מנת להימנע מהצללה (איור 2א).



איור 1. שטח הניסוי בעיר ישי. (א) חלקיות המחקר על גבי צילום אויר של העיר. (ב) מיקום המינים המוניטרים בחלקיות.

המדידות נערכו בתדריות של פעם בחודש, בין אוקטובר 2021 למרץ 2023 (מדידות הפמ"ע והנתונים הספקטראליים מהרףן החלו במרץ 2022) בין השעות 11-13 (טור התamaha עונתית על מנת לתפוס את שיא היום). פוטנציאל המים בעלה (WP) נמדד באמצעות תא לחץ (איור 2ב), כאשר כל עליה שנחנתר הוכנס מיד לשקית ניילון אטומה ומקורה עד למדידה. בנוסף למדידות של פמ"ע נמדד גם פרמטרים של חילוף גזים על מנת להעריך את התగובה הפיזיולוגית של העצים לטיפול במצטומם המשקעים, תוך שימוש באנלייזר גזים נייד (IRGA). המדידות בוצעו על עלים בוגרים ובמקביל למדידות של פמ"ע, באופן שאפשר להעריך את תפקוד העלים באתר ולא פגעה בצמח. נמדדנו שני מדדים עיקריים: Ch - קצב קליטת פחמן דו-חמצני (bihidotot CO_2 ו- O_2 למ"ר לשניה), ו-Tr - קצב דיזוט מהעלים (bihidotot O_2 ו- H_2O למ"ר לשניה). מדדים אלה שימושו להבנת הדינמיקה של חילוף גזים בצמחים בתגובה לתנאי הסביבה ולשינויים במשטר המים.

נתונים ספקטראליים נאספו באמצעות מצלמה היפר-ספקטראלית מסוג Nano-Hyperspec של חברת Headwall Photonics. המצלמה בעלת טווח ספקטראלי של 400-1000 ננומטר (274 ערוצים) מותקנת על גבי רחפן DJI Matrice 600 Pro (איור 2ג), אשר הנתונים נאספו מגובה של 60 מטרים מעל פני הקרקע (הרזולוציה המרחבית המתקבלת היא של כ-2 ס"מ לפיקסל).

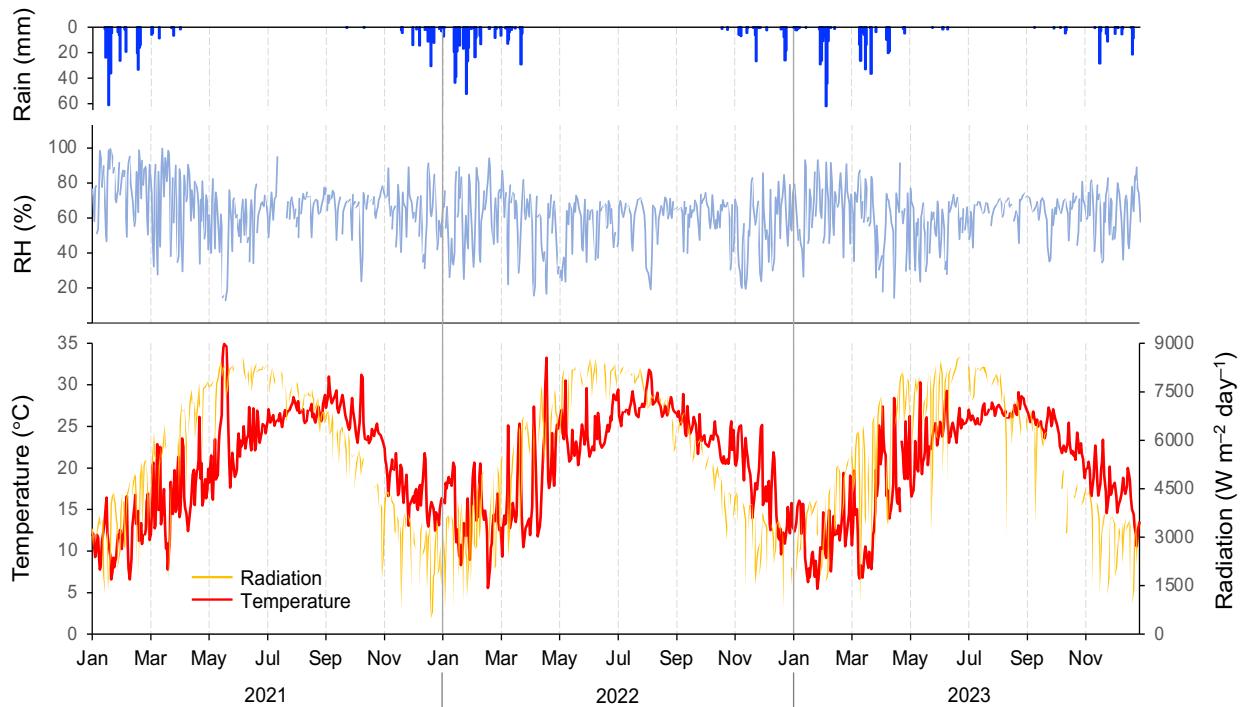


איור 2. (א) מערכת הפחחת הגשם. (ב) מדידת פמ"ע בעזרת תא לחץ. (ג) רחפן רחפן תא לחץ. Nano Hyperspec עם המצלמה ההיפרספקטראלית.

2.1. תוצאות מדידות השדה

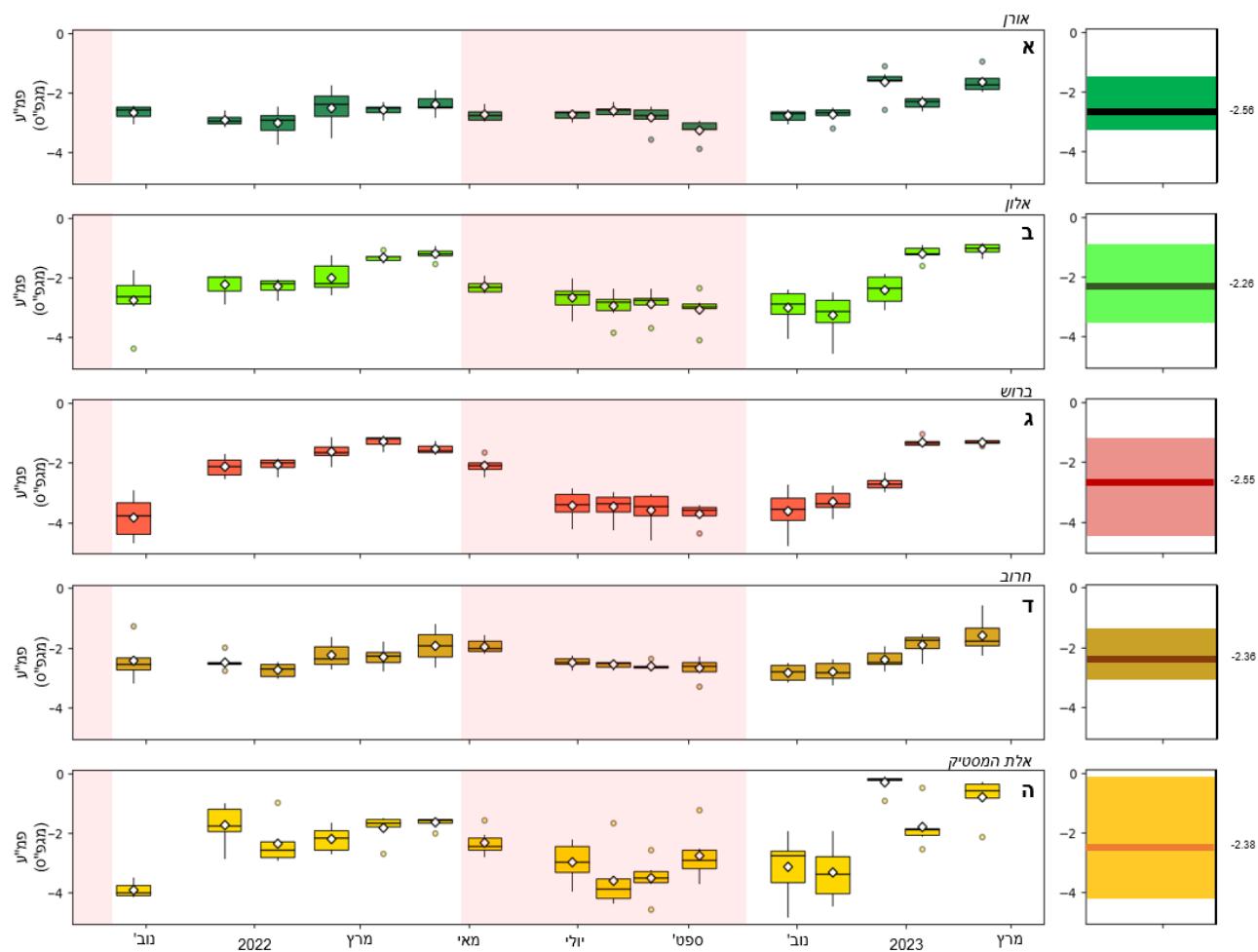
תנאים מטאורולוגיים, מדידות פמ"ע (LWP) , פמ"ג (SWP) ותכולת מים בקרקע (SWC).

חלקוות הניסוי ממוקמות בעיר ישי, הסמור לתחנה המטאורולוגית בית ג'ימאל (כ-2.5 ק"מ מהחקלות), והנתונים נלקחו מאתר השירות המטאורולוגי. **איור 3** מתאר את התנאים המטאורולוגיים לאורך שנות הניסוי (טמפרטורה ממוצעת יומית, קרינה מצטברת יומית, ממוצע יומי של לחות יחסית ומשקעים). ניתן לראות את העונתיות הבוררה, ואת הפיזור של המשקעים לאורך החורף. חורף 22/23 התאפיין בפיזור משמעותי של המשקעים.



איור 3. נתוני קרינה מצטברת, טמפרטורה ממוצעת, לחות יחסית ומשקעים יומיים לאורך חודשים הניסוי. הנתונים נלקחו מתחנת השירות המטאורולוגי בבית ג'ימאל, הקרוובה כ-2.5 ק"מ מחלקות הניסוי.

פמ"ע נמדד בכל המינים החל מאוקטובר 21 ועד ליום 23. התוצאות לאורך שנה וחצי מראות את הדינמיקה העונתית של הפמ"ע, עם ערכים שליליים יותר במהלך הקיץ, בין בחלקות הביקורת ובין בחלוקת הפחחתת הגשם. ממוצעו הפמ"ע לפי מין, המציגים את ההשתנות של הפמ"ע לאורך העונה, מוצגים באירור 4.



איור 4. ממוצעו הפמ"ע לפי מין לאורך חודשי המדידה.

ניתן לראות כי בחלוקת למינים ישנים הבדלים משמעותיים בין השתנות העונתית של הפמ"ע במינים השונים – בעוד האורן והחרוב מציגים עונתיות מתונה ביותר (סביב 2.5 – 2.36 MPa –, בהתאם), ניתן להזיהות עונתיות מודגשת הרבה יותר בברוש ובאלון, והעונתיות הקיצונית ביותר נראית באלת המסתיק, שם הפמ"ע מגיעה עד ל-2.38 MPa – בחודשי הקיץ. תוצאות ניתוח ANOVA על נתונים הפמ"ע מוצגות בטבלה 1.

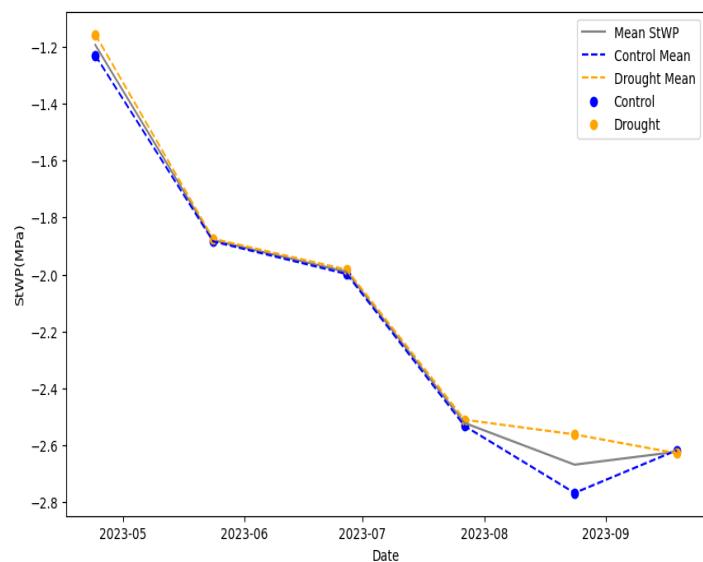
למרות שנראה הבדל מובהק בין הפמ"ע של המינים ובין התאריכים השונים, לא נמצא הבדל מובהק בין הטיפולים. כמובן, שהפחחתת הגשם לא השפיעה באופן שונה על הפמ"ע של המינים השונים.

טבלה 1. ניתוח סטטיסטי ANOVA של השפעת מין (Species), זמן (Time), וטיפול הפחחתת הגוף (Water) על ערכי הפמ"ג (LWP). המובקהות חושבה לאחר בדיקת התפלגות נורמלית של הנתונים בעזרת ניתוח-Shapiro-Wilk. בהדגשה, השפעות מובקהות ברמת $p < 0.05$.

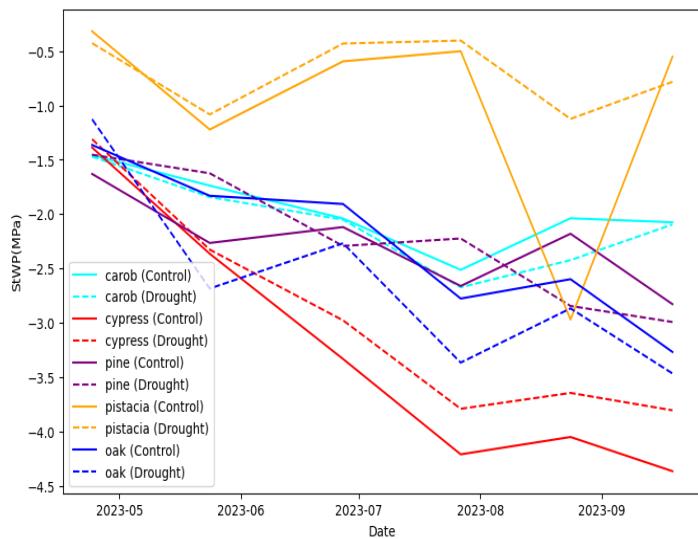
Factor	F Ratio	Prob > F
Species (S)	32.424	<.0001
Time (T)	48.092	<.0001
Water (W)	0.0016	0.9678
S×T	8.4325	<.0001
S×W	3.3437	0.0103
T×W	0.8086	0.7160
S×T×W	0.9398	0.6318

בחודשים אפריל-ספטמבר (מלבד Mai) הוכחשה גם מדידת פוטנציאלי מים בגזע (פמ"ג). ממוצעו ערכי הפמ"ג בחלוקת הביקורת וההפחחתה מוצגים באירור 5, ומהמוצעים לפי המינים השונים באירור 6. תוצאות ניתוח ANOVA מוצגות בטבלה 2. ניתן לראות כי על אף שמדובר בלמעלה משנה וחצי (כולל 2 עונות גשומות) מאז שהותקנה מערכת הפחחתת הגוף, לא ניתן לראות הבדל מובהק בפמ"ג שבין הטיפולים, בין עצמםם ובין באינטראקציה למינים השונים. זאת על אף שמדד הפחמת הפמ"ג נחשבת למדידה יחסית עמידה יותר לשינויים מקומיים מהפמ"ג.

ניתן לסכם כי על אף מערכת הפחחתת הגוף, לא נראה כי המינים בחלוקת ההפחחתה חווים מצב שונה באופן משמעותי מאשר אלו בחלוקת הביקורת. יתרון שהדבר נובע מכך שיש צורך נוספים של הפחחתה על מנת להזנות הבדל משמעותי, או שעל אף מערכת הפחחתת הגוף, חלחול של מים מחלקות סמכות ונגר עלי מטשטשים את האפקט של הפחחתת הגוף על החלקה.



איור 5. ממוצע פמ"ג בחלוקת הביקורת וההפחחתה לאורך חודשים המדידה.

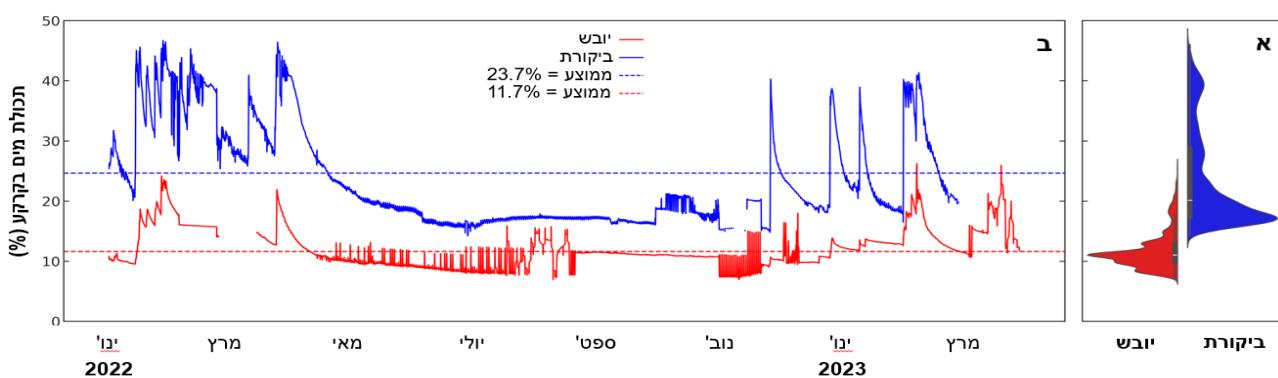


איור 6. ממוצע פמ"ג לפי מין בחלוקת הביקורת וההפחטה לאורך חודשי המדידה.

טבלה 2. ניתוח סטטיסטי ANOVA של השפעת מין (Species), זמן (Time), וטיפול הפחטה הגשם (Water) על ערכי הפמ"ג (StWP). המובאות חושבה לאחר בדיקת התפלגות נורמלית של הנתונים באמצעות ניתוח-Shapiro-Wilk. בהדגשה, השפעות מובאות ברמת $p < 0.05$.

Factor	F Ratio	Prob > F
Species (S)	73.519	<.0001
Time (T)	28.512	<.0001
Water (W)	0.0950	0.7584
S×T	3.7133	<.0001
S×W	1.7503	0.1438
T×W	0.2067	0.9591
S×T×W	1.1291	0.3306

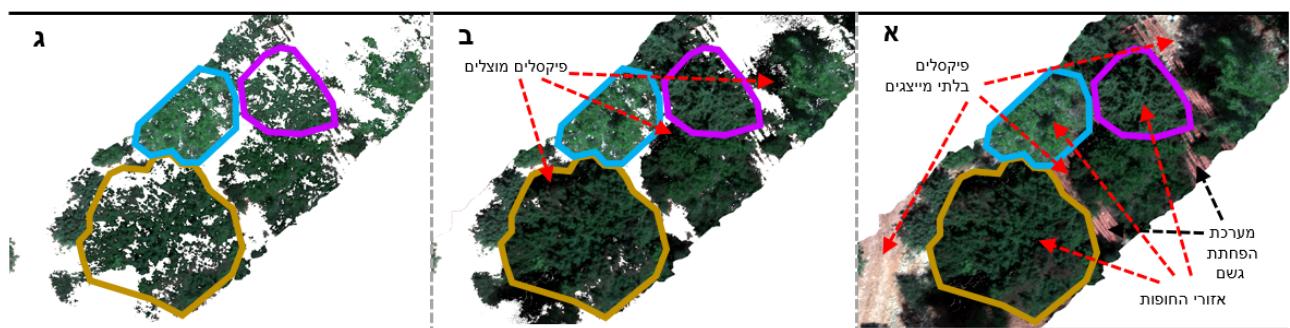
בנוסף לערכי הפמ"ג וההפמ"ג שנמדדו ידנית, בוצע ניטור ריצף של תכולת המים בקרקע באמצעות חיישנים בתת הקרקע, האוגרים ננתונים לאורך השנה. באיור 7 ניתן לראות את תכולת המים הממוצעת בקרקע עבור חלקיota הפחטה הגשם (אדום) ובביקורת (כחול). ניתן לראות באופן ברור שמערכת הפחטה הגשם אכן הביאה לשינויים משמעותיים בתכולת המים בקרקע בעומק של עד 20 ס"מ.



איור 7. תכולת המים בקרקע (%) בטיפולי הפחטה גשם (אדום) ובביקורת (כחול). (א) שינוי תכולת המים בקרקע לאורך זמן בין יוני 2022 למרץ 2023. הקווים המקווקווים מייצגים את הממוצע השנתי לכל טיפול. (ב) התפלגות ערכי תכולת המים לפי טיפול, לאורך כל תקופה המדידה.

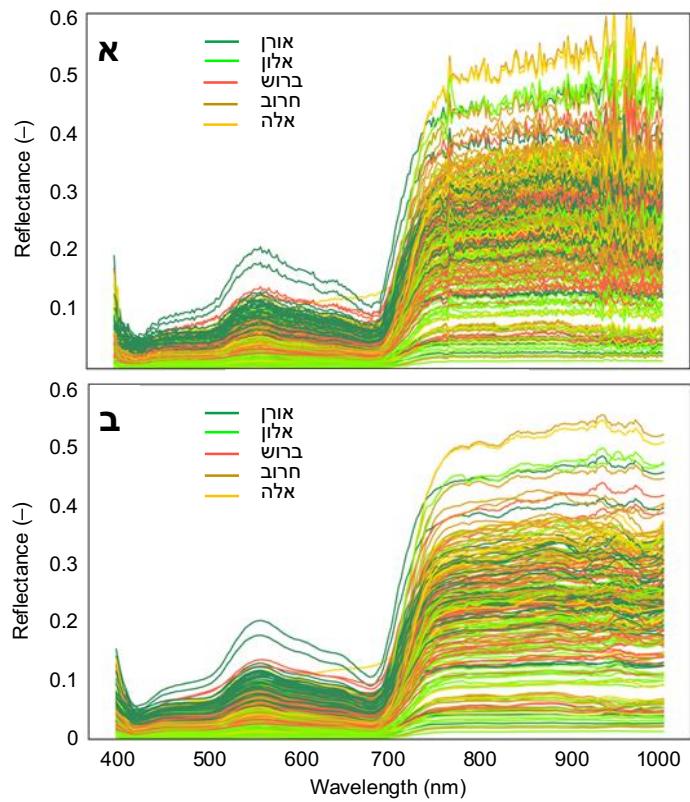
2.2.1. אינדקסים ליניאריים

על מנת לייצר את המדרדים הרלוונטיים לחופות העצים, הנתונים הספקטרליים עברו תהליך עיבוד ראשוני הכלול, עיגון מרחבי, חישוב החזריות באמצעות משטח ייחוס המוצב בשטח, וכן תהליכי של יצירת מזואיקה לתצלום מרחבית רציף (איור 8א). לאחר העיבוד הראשוני, בוצע מיסוך של פיקסלים לא מייצגים בשני שלבים (איור 8). בשלב ראשון, מוסכו פיקסלים שאיןם מייצגים עלווה בעזרת ערך סף של אינדקס NDVI-ND. במקרה שלנו, מצאנו שפיקסלים בעלי ערך NDVI קטן מ-0.3 מייצגים בעיקר קרקע או צומח שעובני מתחת הקרקע (איור 8ב). בשלב שני, מוסכו פיקסלים מוצלים (עמ' צל), המשפיעים על החזריות של העצים. לצורך כך, השתמשנו בעובדה שהחזריות של שטח מוצל מאופיין בערכים נמוכים יחסית של החזריות בעורוצי האינפרא-אדום. גם כאן השתמשנו בערך סף בכדי לנוקוט את התמונה ולהישאר רק עם פיקסלים מייצגים של העצים (איור 8ג).



איור 8. דוגמה למיסוך פיקסלים לא רלוונטיים בדיםות הרחפן. (א) בדיםות נתן לראות את הקרקע, מערכת הפחתת הגשם (צינורות כתומים בתמונה), וצל. (ב) בשלב ראשון, עשינו שימוש בערך סף של $NDVI = 0.3$ כדי להיפטר מפיקסלים של קרקע, שעובנים מתחת-הקרקע ומערכות הפחתת הגשם. (ג) בשלב שני, השתמשנו בערך סף לחזריות בתחום האינפרא-אדום ובידעה שבאזור צל, החזריות בתחום הנ"ל קטנה משמעותית מבאזורים מוארים.

לאחר שנותרנו רק עם הפיקסלים שמייצגים את חופות העצים, מתבצע מיצוע של ערכי החזריות של כל הפיקסלים לכל אחד מהעורוצים שהמצלמה צולמה בהם (כאמור, 274 ערוצים בתחום 400-1000 ננומטר), לקבלת חתימה ספקטרלית המייצגת את החזריות בכלל הפיקסלים של חוף העץ, לאורך כל 274 הערוצים. החתימה עברה חילקה בשיטת "Savizki-Golay filter", אחת השיטות המקובלות ביותר לה החלקה של נתונים ספקטרליים. ההבדל בין הנתונים לפני ואחרי החלוקת מוצג באיור 9.



איור 9. החתימה הספקטרלית של העצים השונים (א) לפני (ב) ואחרי חילקה בשיטת סביצקי-גולאי.

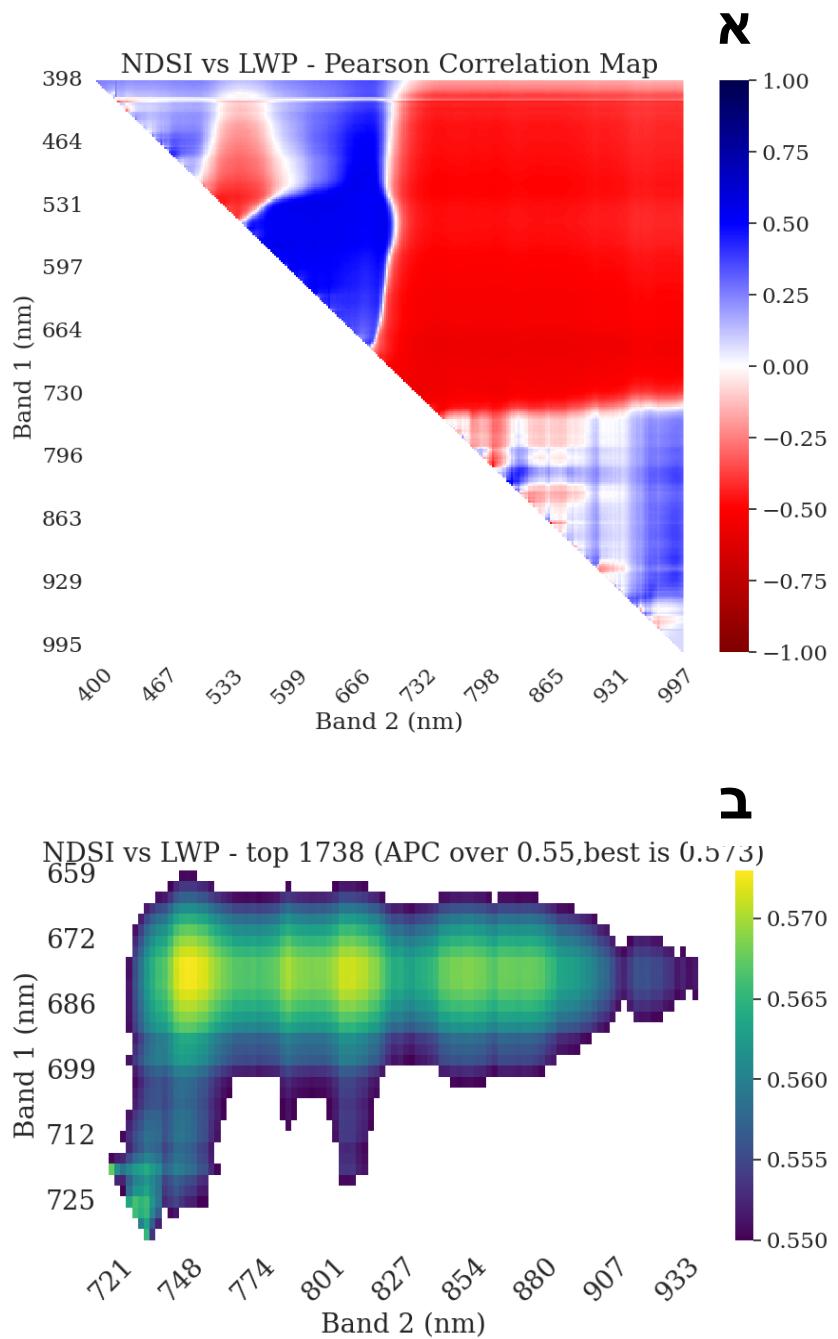
לאחר חילקת הנתונים, נבנתו 12 מדדים ספקטרליים מוכרים המשמשים למידול צמחייה, מים ופמ"ע על מנת לזהות מי מהם מייצר מטען משמעותי לשינוי בפמ"ע. רשימת המדדים והנוסחאות שלהם מוצגות בטבלה 3. בנוסף, נועה ניתוח מפת הקורלציות של ערכי ההחזריות של כלל העוזרים, עבור החופות שנמדדו. המטרה היא למצוא את המדד הספקטרלי מסווג NDSI שיציג את ההתאמה הטובה ביותר בין ערכי המדד לנוטוי הפמ"ע. מדדי NDS מחושבים מערבי ההחזריות של שני עוזרים, לפי הנוסחה הבאה:

$$NDSI = \frac{Band\ 1 - Band\ 2}{Band\ 1 + Band\ 2}$$

טבלה 3. מדדים ספקטרליים מוכרים המשמשים לצמחייה, מים ופמ"ע.

Index	Name	Formula
NDVI	Normalized difference vegetation index	$\frac{\rho_{802} - \rho_{670}}{\rho_{802} + \rho_{670}}$
PRI	Photochemical reflectance index	$\frac{\rho_{531} - \rho_{570}}{\rho_{531} + \rho_{570}}$
940/960	Reflectance ratio of 940 and 960 nm	$\frac{\rho_{940}}{\rho_{960}}$
940/960/ NDVI	Reflectance ratio of 940/960 nm and NDVI	$\frac{\rho_{940}}{\rho_{960}} / \text{NDVI}$
EVI2	Enhanced vegetation index 2 nd version	$2.5 \left(\frac{\rho_{802} - \rho_{670}}{\rho_{802} + (6 \rho_{670}) - (7.5 \rho_{480}) + 1} \right)$
COSBNDI	Combined overtone of stretching bands – normalized difference index	$\frac{\rho_{660} - \rho_{420}}{\rho_{660} + \rho_{420}}$
FOSBNDI	Forth overtone of stretching bands – normalized difference index	$\frac{\rho_{529} - \rho_{698}}{\rho_{529} + \rho_{698}}$
SAPSBNDI	Small absorption peak of stretching bands – normalized difference index	$\frac{\rho_{750} - \rho_{970}}{\rho_{750} + \rho_{970}}$
WASCOSBNDI	Water absorption shoulder due to combined overtone of stretching bands – normalized difference index	$\frac{\rho_{800} - \rho_{847}}{\rho_{800} + \rho_{847}}$
NDWSI	Normalized different water stress index	$\frac{\rho_{850} - \rho_{970}}{\rho_{850} + \rho_{970}}$
MNDWI	The modified normalized difference water index	$\frac{\rho_{546} - \rho_{802}}{\rho_{546} + \rho_{802}}$
WI	Water index	$\frac{\rho_{970}}{\rho_{900}}$

בעזרת קוד מתאים, מופקת מפת קורלציות פירסון המשקפת את טיב ההתאמה בין סט ערכי הפמ"ע לבין סטיהם של ערכי NDVI המופקים מכל זוג אפשרי של ערוצים. באIOR 10 ניתן לראות שישנו במפת הקורלציות אזורים שונים, עם מתאם פירסון בין ה-NDVI לפמ"ע חיובי או שלילי והבדל בערכים עצם. בעזרת מפת הקורלציות בחרנו את זוגות הערוצים שה-NDVI שלהם נמצא בהתאם גובה עם פמ"ע. ערוצים אלו נכנסו למודל של מידת המכונה. תחילה נעשה עבור הסט המלא של כל נתוני המחקר, וגם עבור חלקים ממנו, או ממוצעים – כגון ממוצע חלקות וכדומה. זה נושא לממצוא את הערוצים המצליחים ביותר נעשה סינון ראשוני, כך שמדובר רק זוגות שלהם מתאימים גובה דיו על מנת למצוא את הערוצים המצליחים ביותר נעשה סינון ראשוני, כך שמדובר רק זוגות שלהם מתאימים גובה דיו (מתאם פירסון > 0.55). אIOR 10 ב מציג את איזור מתאם פירסון הגובה מ-0.55 בחלוקת הקורלציה עבור כל העצים שנמדדו בכל התאריכים.



איור 10. מפות קורלציה בין NDSI ולפמ"ע, עבור כל העצים בכלל התאריכים. (א) מפת קורלציה לפי מתאם פירסון, ניתן להבדיל בין מתאם חיובי (כחול) לשיליי (אדום). (ב) מפת קורלציות לפי ערך מוחלט של מתאם פירסון, המכילה רק את הקורלציות עם מתאם הגובה מ-0.55.

2.2.2. מודלי למידת מכונה

כאמור, חישבו אינדקס NDSI עבור כל זוג ערוצים אפשרי, מה שהניב למעלה 37,000 צירופים. כדי למנוע חפיפה מיותרת בין תחומי גל סמוכים (שעלולה לגרום לכפיליות), הוסרו אינדקסים המבוססים על תחומים מרוחקים פחות מ-10 ננומטר זה מזה. לאחר מכן נבחרו חמישה האינדקסים שהראו את הקורלציה הגבוהה ביותר עם פמ"ע - האינדקס המוביל, וארבעה נוספים שהציגו ערכי شيئا מוקמיים עבור כלל המינים.

האינדקסים הספקטרליים המציגים את ההתאמנה הגבוהה ביותר עם פמ"ע (על פי סוף מוגדר מראש) בסט האימון נוצלו כתוכנות קלט למודלים של מיידת מכונה.

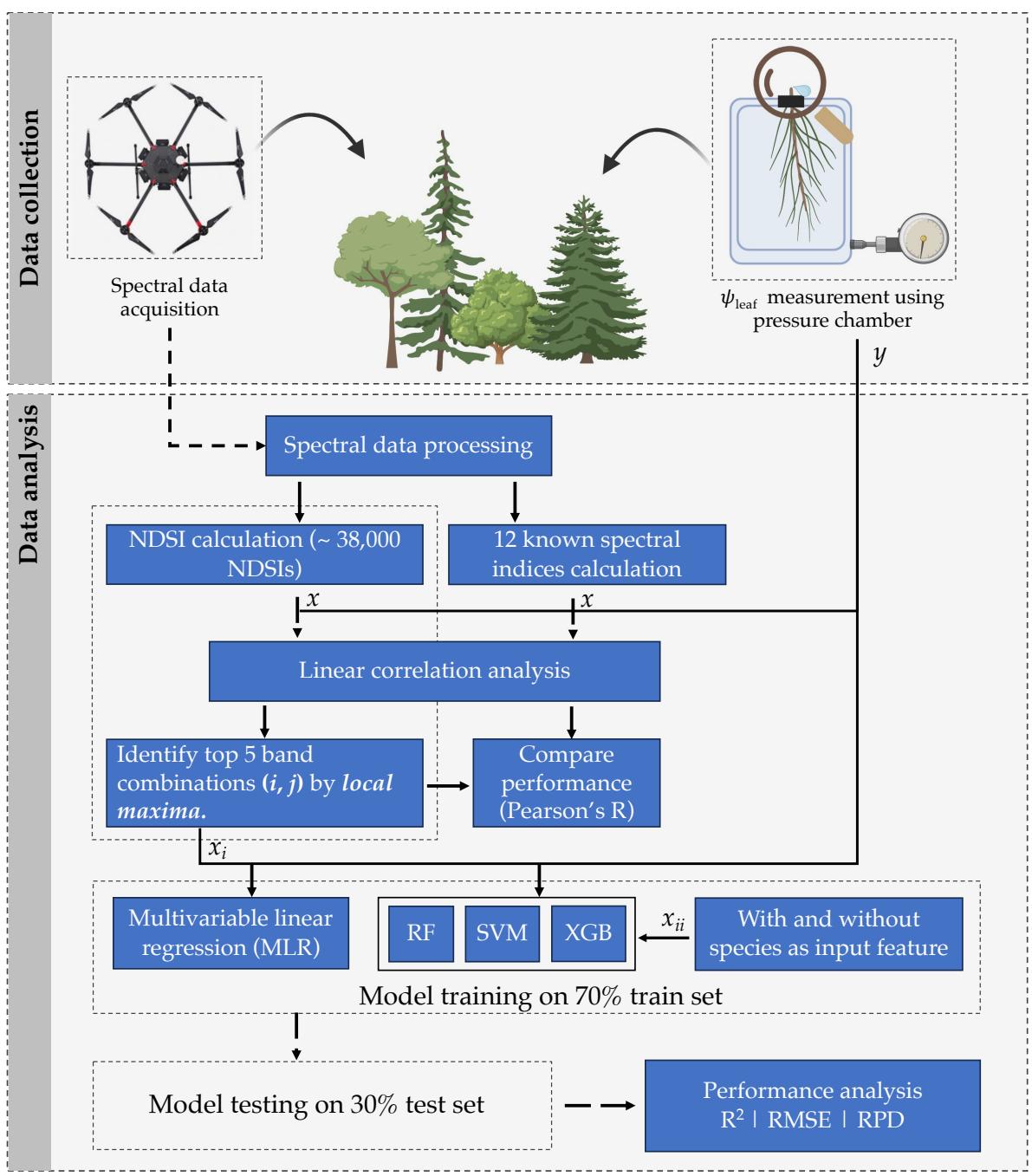
השתמשנו במודל רגסיה לינארית (**MLR**) ובשלושה אלגוריתמים מבוססי למידת מכונה לא לינאריים:

RF: שיטת למידה קולקטיבית שבונה מספר עצי החלטה שונים ומצרפת את הניחושים שלהם, תוך התחשבות בשונות והימנעות של תאימות יתר (Breiman 2001).

XGB : מערכת שיפור מבוססת עץ נפוצה, שבונה אנסמבל של מודלי ניחוש חלשים בצורה רציפה, כאשר כל מודל מצליח מתקן את השגיאות של קודמו (Chen and Guestrin 2016).

SVM (Support Vector Machine): אלגוריתם למידה מונחית, שבונה על-מיישורים במרחב רב ממדים כדי למקסם את הרוחות בין מחלקות, המסוגל לטפל בקשרים לא לינאריים באמצעות פונקציות ליבה (Cortes and Vapnik 1995).

המודלים הוכשרו על 70% מהנתונים (סט האימון) באמצעות שילוב של חיפוש היפר-פרמטרים אקראי ואימות מובלב (cross validation). ביצועים הוערכו על ה-30% הנתרים (סט המבחן) באמצעות מדדים כגון מקדם המתאם (R^2), שגיאת הריבוע השורשית הממוצעת (RMSE), יחס ביצועים לשגיאה (RPD) ושגיאה ממוצעת מוחלטת (MAE). בנוסף, נוצר מודל אנסמבל (AVG) על ידי מיצוע הניחושים מכל המודלים הפרטיים. נערכו ניתוח חשיבות התכונות כדי לזהות את האינדקסים הספקטרליים המשפיעים ביותר עבור כל מודל. בשלב הבא, מידע על מין העצים הוכנס כמשתנה קטגורית באמצעות קידוד-hot-echo ונוסף לתוכנות הקלט. תהליכי המידול חוזרים על עצמו וכלל תיקון של היפר-פרמטרים והערכת ביצועים על סט המבחן. בנוסף, חושבו RMSE ו- R^2 - נפרדים עבור כל מין עץ כדי להעריך את ביצועי המודל על כל המינים השונים. איור 11 מציג את סכמת המודלים.



איור 11. תרשيم סכמטי של מסגרת המחקר – משלבי איסוף הנתונים ועד בניית מודלים של למידת מכונה. א' ו' י מייצגים את המשתנים המסביר והמוסבר (בהתאמה). א' ו' י מייצגים את קבוצות המשתנים המסבירים (predictors) שהתבססו על חמישה אינדקסי ה-NDSI המובילים, ללא (י) ועם (י) מידע על מין הצמח כמשתנה קלט. הקיצורים RF, SVM, XGB מציינים את האלגוריתמים Random Forest, Support Vector Machine ו- Extreme Gradient Boosting, בהתאמה.

2.3. תוצאות ניתוח מודלים

מבין 12 האינדקסים המוכרים שנבחנו, NDVI הראה את הקשר הלינארי החזק ביותר עם פמ"ע עבור עצ הברוש עם מתאם פירסון של 0.86. שאר האינדקסים הציגו קשרים ביןוניים או חלשים בלבד. חלקם נמצאו בקורסציה חיובית עם פמ"ע ואחרים בשלילית (ראו **טבלה 4**).

טבלה 4. מתאם פירסון בין המדדים הספקטרליים השונים לבין ערכי הפמ"ע.

Index	Pine	Oak	Cypress	Carob	Pistacia	All
NDVI	0.56	0.70	0.86	0.53	0.54	0.57
PRI	0.31	-0.18	0.19	-0.18	0.55	0.09
940/960	0.20	0.24	0.31	0.29	0.48	0.29
940/960 / NDVI	-0.51	-0.65	-0.84	-0.38	-0.34	-0.48
EVI2	0.23	0.10	-0.08	-0.06	0.24	0.08
COSBNDI	-0.32	-0.13	-0.39	-0.18	-0.63	-0.29
FOSBNDI	0.41	0.38	0.63	0.21	0.50	0.43
SAPSBNDI	0.11	0.14	0.49	0.26	0.60	0.27
WASCOSBNDI	-0.22	-0.16	0.26	0.03	0.13	0.03
NDWSI	0.15	0.21	0.50	0.27	0.60	0.29
NDWI	-0.52	-0.65	-0.83	-0.51	-0.67	-0.48
WI	-0.21	-0.27	-0.48	-0.26	-0.60	-0.31
NDSI (680/750)	0.55	0.69	0.86	0.54	0.60	0.57

גם אינדקס NDSI המוצלח ביותר שנמצא, המבוסס על ערכים 750 ו-680 ננומטר, לא שיפר את הקשר הליניארי עם פמ"ע. כאשר נבחנו כל המינים יחד, גם האינדקסים הטובים ביותר (NDVI וה-NDSI המוביל) הציגו מתאם ביןוני בלבד, עם ערך של $R = 0.57$.

מודל הרגרסיה הליניארית (**MLR**) נתן תוצאות רק מעט יותר טובות מאשר מדדים ספקטרליים (סוג NDSI) אחרים ($R^2=0.40$ לועמת $R^2=0.34$ ב-NDSI). המודל דרש 9 ערכים ו-5 זוגות כדי לייצר את מודל הרגרסיה הליניארית הטוב ביותר (טבלה 5).

טבלה 5. מודל רגרסיה ליניארית עם צירופי NDSI הטוביים ביותר.

NDSI (band1/band2) combinations	R^2	RMSE
530/623	0.31	0.74
530/623, 680/750	0.35	0.72
530/623, 680/750, 660/940	0.35	0.72
530/623, 680/750, 660/940, 519/750	0.35	0.72
530/623, 680/750, 660/940, 519/750, 605/709	0.40	0.70

המודלים הלא ליניארים (RF,SVM ו-XGB) לא הציגו תוצאות טובות יותר (טבלה 6), אבל הביצועים השתפרו כאשר המינים הוכנסו כתכונה במודל (טבלה 7). המודל הטוב ביותר היה SVM, עם R^2 של 0.53 ו- RPD של 1.47.

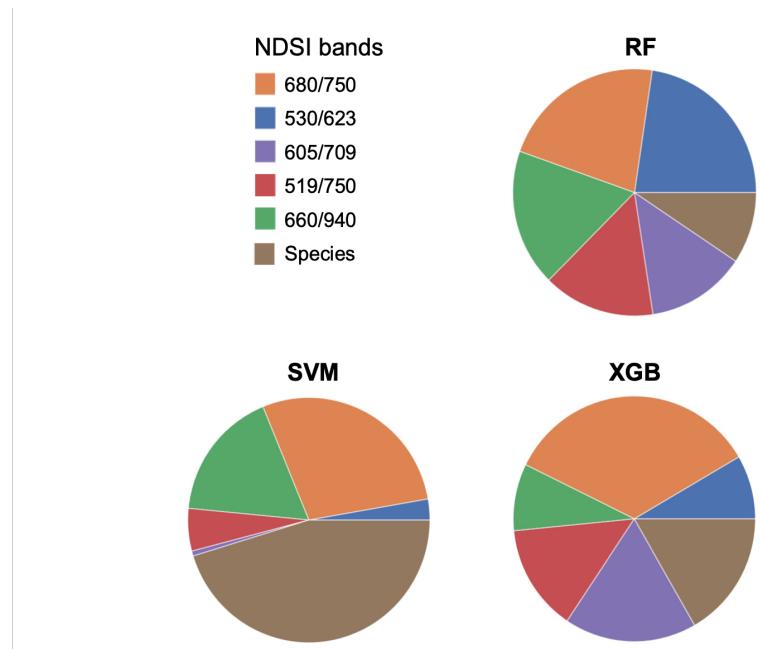
טבלה 6. סטטיסטיות עבור מודלי למידת המכונה השונים - ללא התייחסות למינים.

Model	R^2	MAE	RMSE	RPD
RF	0.40	0.59	0.74	1.30
SVM	0.35	0.58	0.77	1.25
XGB	0.40	0.58	0.74	1.30
<i>Averaged model</i>	0.41	0.58	0.74	1.31

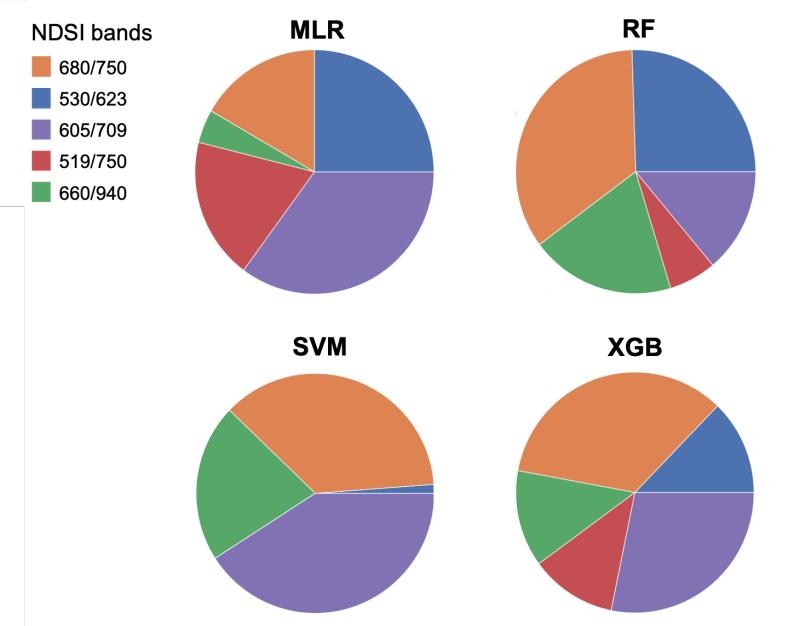
טבלה 7. סטטיסטיות עבור מודלי למידת המכונה השונים – כאשר המינים נחשבים כתוכנה.

Model	R^2	MAE	RMSE	RPD
RF	0.42	0.59	0.74	1.32
SVM	0.53	0.50	0.67	1.47
XGB	0.47	0.57	0.71	1.38
<i>Averaged model</i>	0.52	0.52	0.67	1.46

כל ארבעת המודלים הציגו את הביצועים הטובים ביותר עם NDSI של אותם צירופי עrazois. אולם החשיבות של כל צירוף NDSI במודל השתנתה בין המקירה בו המודל השתמש במינים כתוכנה (איור 12) או לא השתמש במינים כתוכנה (איור 13).

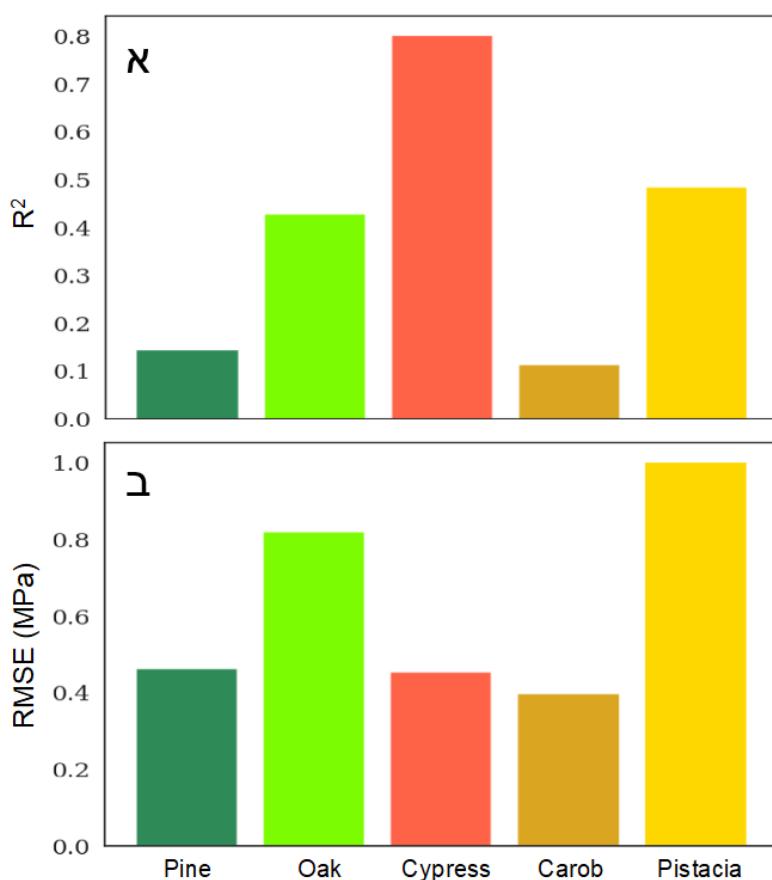


איור 12. גרף עוגה של חשיבות התכונות של צירופי העrazois ב-NDSI השונים – התייחסות למין כתוכנה.



איור 13. גרף עוגה של חשיבות התכונות של צירופי העורוצים ב-NDVI השונים – ללא התייחסות למין כתכמה.

לבסוף, המודל הטוב ביותר (SVM תוך שימוש במינימום כתכמה) חזה את הפמ"ע של הברש והאלה הרבה יותר טוב מאשר שאר המינים, ובמיוחד האורן והחרוב (איור 14).



איור 14. ביצועי מודל SVM לפי המינים השונים: (א) ערכי R^2 . (ב) ערכי RMSE.

איור 15 מסכם את המודל הטוב ביותר (והשני הטוב ביותר) עבור כל מין. ניכר שהמודלים פשוטים המשמשים במדד סטטיסטי ייחד או בזוג מדי NDSI יכולים לחזות פמ"ע בדיק ביןוני, וכי המודלים של למידת מכונה לא עלו על ביצועיהם באופן משמעותי.

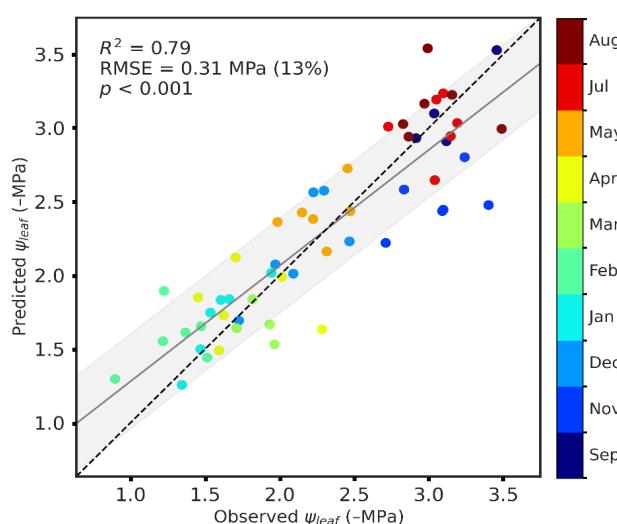
	Pine	Oak	Cypress	Carob	Pistacia	All
Best model	NDSI (909/685)	NDSI (909/685)	SVM (w/ species)	NDSI (816/712)	SVM (w/ species)	SVM (w/ species)
R ²	0.33	0.50	0.80	0.28	0.49	0.53
2nd best model	NDVI	NDVI	NDSI (816/712)	NDVI	NDSI (816/712)	Avg ML (w/o species)
R ²	0.31	0.48	0.78	0.28	0.47	0.41
Difference	0.02	0.02	0.02	0.00	0.02	0.12

■ ML model
■ Spectral index

איור 15. סיכום המודלים הטובים ביותר לפי המינים השונים, ועבור כולם יחד.

בשל ביצועים מתונים במודלים שהתמקדו במינים בודדים, נוסה גם מודל כללי שהתבסס על ממוצע מדידות ברמת החלקה שכיל את כל חמישה המינים יחד. המודל הכללי שהתבסס על NDSI המשלב תחומי גל 816 ו-712 ננומטר נמצא כמדד החיזוי הטוב ביותר עבור פמ"ע בrama זו.

איור 16 מציג את הקשר בין הנטה לחיזוי במודל הכללי, עם $R^2 = 0.79$ ו- RMSE יחסית של 13%, תוצאה מובהקת סטטיסטית ($p < 0.001$).



איור 16. השוואה בין ערכים חווים ונמדדים של פמ"ע עבור נתוני ממוצע ברמת חלקה. כל נקודה מייצגת את הממוצע של כל המינים בחלוקת בתאריך דגימה אחד. הצבעים מצינים את חודש המדידה/תחזית. הקו המקווקו מייצג קו 1:1.

2.4. מודל סיוג

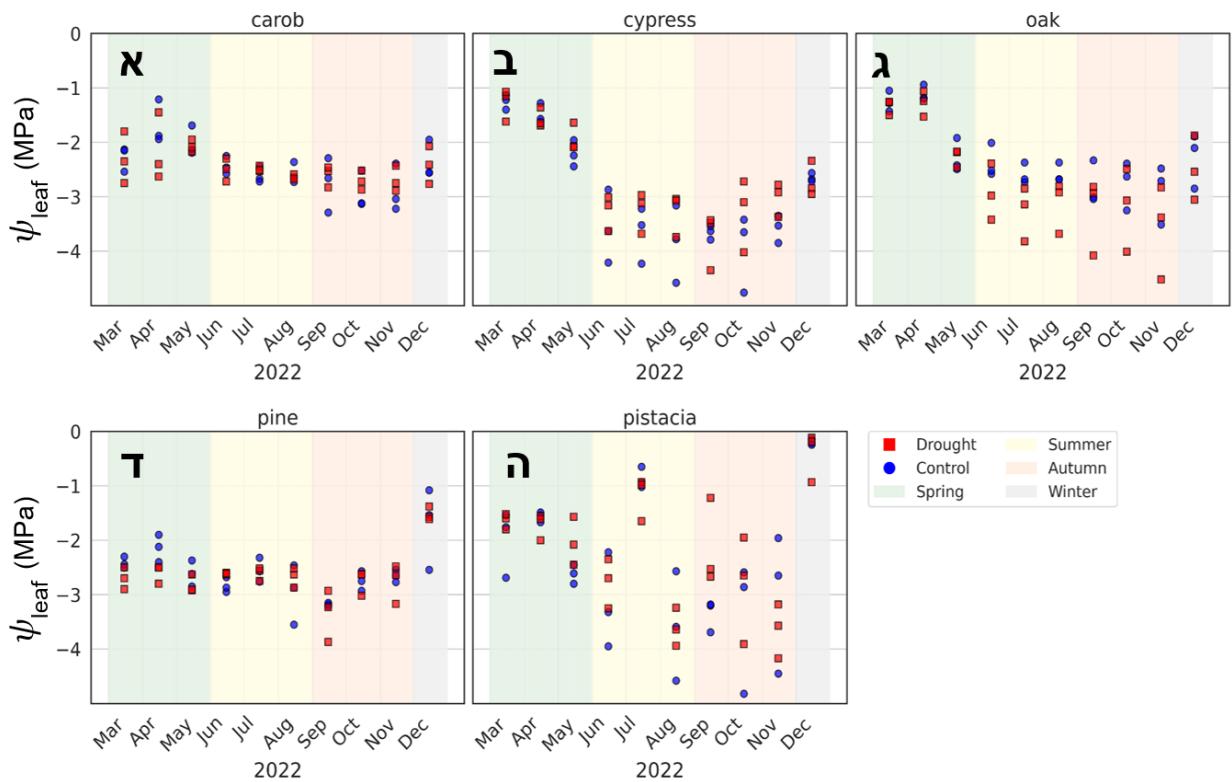
2.4.1. שיטות

בשלב זה, על מנת לפתח כל יישומי המאפשר המלצות ממשקיות, נותחו הנתונים הספקטרליים שהופקו מהרחפן מחדש, והפעם בהתאם למאפיינים של 3 מערכות לווייניות - *Sentinel-2*, *VENµS* ו-*RF*. יצרנו מתוך נתונים הרחפן ערוצים סינטטיים, על ידי מציאת הנתונים באורכי הגל המקבילים לערכאים שקיים בלוויינים (בתוך הטווח הספקטרלי שבו עבד החישון של הרחפן – 400-1000 ננומטר). לאחר אימון שלושה מודלי למידת מכונה (*RF*, *WMS*-*LR*) על הנתונים הסינטטיים, הופעל המודל שהגיע לתוצאות הטובות ביותר (*WMS*) על נתונים לוין אמיתיים של שלושה ערוצות במورد מפל הגשם בישראל, כדי לבדוק את יכולת המודל לזהות מוקדם של עומדי יער בעקבות יובש. לצורך בוחינת המודל הורדו נתונים לוין *VENµS*, עברו עומדים בעירות קק"ל גלבוע, צרעה ושהירה. הורדו תמונות לוין מחודשים זהים לאורך מספר שנים, על מנת להימנע משנות על רקע עונתי. צלומי הלווין מנטרים שלושה שלבים – התנאים לפני הצורת, שנת בצורת, ומצב העיר בשנה העוקבת (כולל פעולות ממשק שבוצעו ע"י קק"ל).

בשלב הראשון, הופלו המודלים על מלאו 247 הערוצים שמספקת מצלמת הרחפן, על 70% מהנתונים (sett האימון). בשלב הבא היה צורך לבחור פחות ערוצים, על מנת להימנע מרוב מדדיות והתאמת יתר (overfitting) של המודל. עבור כל אחד מהמודלים נבחרו רק הערוצים שאחראים על 80% מהיכולת של המודל לסוג את הנתונים, ורק ערוצים רחוקים אחד מהשני לפחות 10 ננומטר. לאחר מכן, נבחן המודל על ה-30% של הנתונים (sett המבחן), ונערך ניתוח סטטיסטי מפורט על מנת להשוות בין יכולות הסיווג של שלושת המודלים. בעקבות התוצאה, נבחר מודל *WMS* על גבי לוין *VENµS*.

2.4.2. תוצאות

בחינת ערכי הפמ"ע בחלוקת היובש והביקורת לאורך העונות השונות מראה כי הפחיתה כמות המים הנכנסים (כמעט בחצי) לא השפיעה באופן מובהק על פמ"ע ברוב המינים ועונות השנה (איך 17). החיריגים היו עצם, שהראהו ערכים שליליים יותר בקי"ז ובסטוי, עצם האורן, שהציג ערכים נמוכים יותר באביב.



איור 17. מדידות פמ"ע בחמישה מיני עצים, תחת תנאי ביקורת וטיפול בהפחיתה גשם. תרשימי פיזור מציגים פמ"ע ביחסות של מגה-פסקל עבורי: (א) חרוב, (ב) ברוש, (ג) אלון, (ד) אורן, ו-(ה) אלה, במהלך מרץ עד דצמבר 2022. כל נקודה מייצגת מדידה של עץ בודד; עצים בטיפול ביקורת מסוימים בעיגולים כחולים, ועצים בטיפול בוצמצום משקעים – בריבועים אדומים. צבע הרקע מציין את העונה: אביב (ירוק; מרץ–מאי), קיץ (צהוב; יוני–אוגוסט), סתיו (אדום בהיר; ספטember–נובמבר), וחורף (אפור; דצמבר–פברואר).

בדיקות סטטיסטיות מסוג *t-test* אישרו כי למרות הפחיתה משמעותית בכמות המים, לא נמצא השפעה ברורה על הפיזיולוגיה של רוב העצים – לפחות לא במידדים של פמ"ע, קצב קליטת פד"ח (A_n) ודיאת – (T_r) לאורך רוב העונה ובמרבית המינים (טבלה 8).

טבלה 8. תוצאות מבחני τ דו-צדדיים להשוואת השפעת טיפול הביקורת והיבוש על פמ"ע (ψ_{leaf} , דיות, T_r) וקליטת פחמן (A_n) בחמישה מיני עצים לאורך ארבע עונות. ערכי ה- τ מוצגים בטבלה, כאשר השפעות מובהקות (p) מודגשות, והשפעות מובהקות גבולית ($p > 0.1$) מסומנות באותיות נטויות (Italic).

Species	Season	ψ_{leaf} (MPa)	T_r (mmol m ⁻² s ⁻¹)	A_n (μmol m ⁻² s ⁻¹)
Cypress	Spring	0.771	0.994	0.275
	Summer	0.070	0.910	0.542
	Autumn	0.127	0.366	0.335
	Winter	0.770	0.603	0.160
Oak	Spring	0.751	0.127	0.428
	Summer	0.002	0.026	0.101
	Autumn	0.066	0.133	0.472
	Winter	0.666	0.117	0.309
Pine	Spring	0.010	0.249	0.927
	Summer	0.247	0.702	0.274
	Autumn	0.589	0.797	0.577
	Winter	0.672	0.904	0.164
Carob	Spring	0.314	0.336	0.021
	Summer	0.963	0.089	0.117
	Autumn	0.212	0.472	0.338
	Winter	0.841	0.040	0.643
Pistacia	Spring	0.239	0.348	0.754
	Summer	0.978	0.628	0.635
	Autumn	0.381	0.989	0.294
	Winter	0.459	0.454	0.376

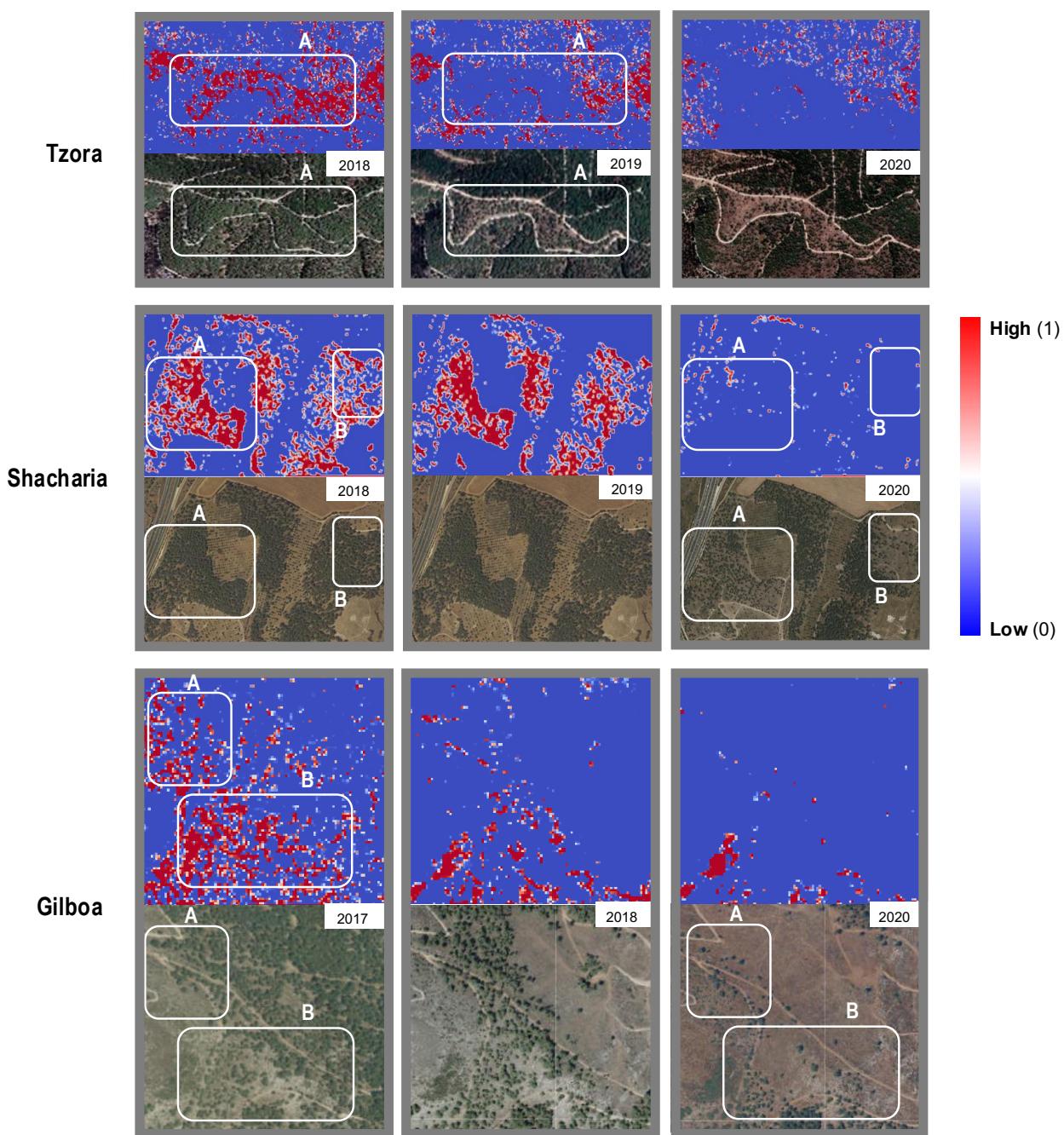
המודל שהציג את הביצועים הטובים ביותר עבור הנתונים הסינטטיים מהלוין היה מודל ה- SVM (טבלה 9). התוצאות הסטטיסטיות מצביעות על כך ש-SVM השיג ביצועים טובים יותר, כמעט בכל המדרדים, בערוצים של כל שלושת הלוינים שנבדקו. מתוך שלושת הלוינים, לוין VENμS הציג את הדיק הגבוה ביותר – ערך דיוק של 0.73 (accuracy) וערך F1 של 0.74 (accuracy).

טבלה 9. מדדי ביצוע של שלושת מודלי למידת המכונה עבור סיווג ביןארי, בהתבסס על נתוני סימולציה ספקטרליים מלויני PlanetScope, VENμS ו-Sentinel-2. הציון הגבוה ביותר עבור כל לויין וממד מודגש.

Satellite	PlanetScope			VENμS			Sentinel-2		
	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF
Accuracy	0.55	0.65	0.61	0.64	0.74	0.54	0.54	0.70	0.51
Precision	0.51	0.63	0.58	0.58	0.70	0.50	0.50	0.66	0.47
Recall	0.68	0.67	0.56	0.74	0.76	0.41	0.62	0.74	0.44
F1 Score	0.58	0.65	0.57	0.65	0.73	0.45	0.55	0.69	0.45

לא נמצא הבדלים משמעותיים בחשיבות של העורוצים השונים, אם כי התחום B12 (900–920 ננומטר) נמצא כחשוב ביותר (10%), ואילו B2 (400–440 ננומטר) היה הפחות חשוב (7%) ראו איור משלים.(S3).

שימוש במודל SVM עם נתונים אמיתיים מלאוין *SuNEN* בדצמבר 2017 זיהה אזוריים ממוקדים בכל שלושת האתרים שנבדקו, בהם קיימת סבירות גבוהה לעקמת יובש (מסומנים באדום באירור 18). לעומת זאת, תמונות RGB מהאתר הראה לא הרואו סימנים לירידה במצב העיר או חלקות ספציפיות המזויות בסיכון.



איור 18. יישום מודל זיהוי עקמת יובש (SVM) על דימות לווין *SuNEN* בשלושה אתרים עיר ים-תיכוניים לאורך מסוף שנים. כל שורה מייצגת אתר עיר אחר: צרעה (למעלה), שחריה (באמצע), וgilbo (למטה). כל עמודה מציגה שנה שונה (מצינת בפינה הימנית העליונה), עם זוג תמונות: התמונה העליונה מציגה את תחזית הסיכון לעקמת יובש על פי המודל (כחול = סיכון נמוך, אדום = סיכון גבוה), והתמונה התחתונה – את דימות ה-RGB-RGB המתאימים מהלוין. המסגרות הלבנות (A ו-B) מסמנים אזוריים בהם המודל זיהה סיכון גבוה לעקקה (פיקסלים אדומים) עוד לפני שההתקן הפיזי היה גלוי לעין. סרגל הצבעים מימין מיצג את ההסתברות לעקמת יובש, מ-0 (כחול, נמוך) ועד 1 (גובה, אדום).

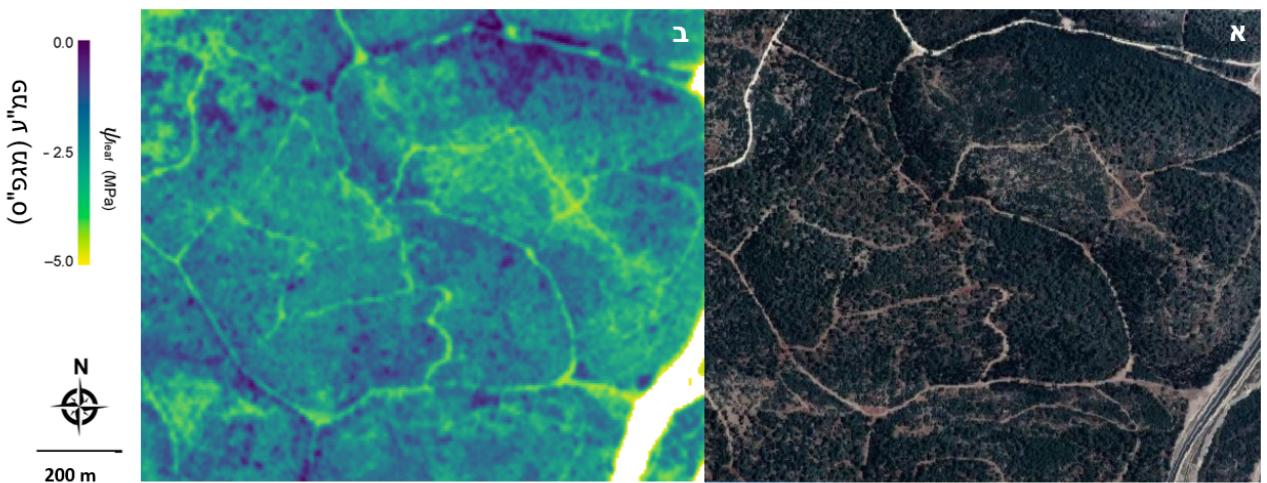
למרות זאת, באתר צרעה דוחה על תמונות עצים חמורה במהלך השנה שלאחר מכן (כפי שנראה בצלום אויר וכפי שדווח על ידי קק"ל), כמעט בדיקו אותו אזור שנחזה מראש בסיכון גבוה על ידי מודל SVM. תצפית דומה חוזרת על עצמה שנתיים לאחר מכן גם באתרם שחריה וgilbo.ע.

מעניין לציין כי ברגע שהנזק נראה לעין (בתמונות ה-RGB), המודל כבר לא זיהה את האזור כמצוי בסיכון לעקבתו יובש (איור 18).

3. מסקנות, המלצות ממשקיות ופערן ידע

למרות שנתיים של הפחתת גשם, נראה כי לא ניתן לזהות הבדל מובהק בערכי הפמ"ע בחלוקת הביקורת וההפקחה. אולם, הבדל זה כן ניתן להיזהו בתוכנות המים בקרקע.

למרות שלא נמדדה עקה שנוצרה על ידי הפחתת הגשם, בעזרת שימוש במידדי NDSI וכן בעזרה באמצעות מודלי למידת מכונה, הצליחנו ליצור מתאם, ועל ידי כך מודל, בין המידע הספקטורי לבין ערכי הפמ"ע בעצים. מודלי למידת המכונה (ובמיוחד מודל SVM) הצליחו להגיע לתוצאות משמעותיות במידול פמ"ע לפי מידע ספקטורי. בנוסף, על ידי מיצוע של ערכי HSND של העצים השונים בזוג הערכים 816/712, ניתן להגיע למידול ברמת דיוק גבואה של הפמ"ע המוצע של החלקה (איור 16). מידול מסוג זה הוא שימושי, כיוון שניתן ליישם אותו בקלות יחסית בעזרת צילומי לווין, שהם בעלי רזולוציה מרחבית נמוכה – ככלומר ממצאים בכל מקרה את ההחזר ממספר חופות עצים בכל פיקסל. דוגמא ליישום של המודל על עיר ישע ניתן לראות באיור 19.



איור 19. חיזוי פמ"ע באמצעות יישום ה"מודל הכללי" על בסיס דימות לווין S-ENVEN בעיר ישע. כבישים מסומנים בלבן. (א) צילום RGB של העיר. (ב) מפת פמ"ע (ביחידות מג'ס).

למרות שלא ניתן היה לזהות את השפעת הפחתת הגשם על הפמ"ע, נראה מתוצאות מודל הסיווג שניתן לזהות בעזרה שימושUrוצים ספקטוריים ברמת דיוק גבואה איזו חלקה עברה הפחתת גשם ואיזו לא, מה שמלמד שהעצים אכן מגיבים להתיישבות הקרקע, גם אם יש להם מגגוננים שמאיפים על החוסר כך שהוא לא מתבטא בפמ"ע. תוכנה זו אפשרה לנו ליצור מודל סיווג, שמנῆה את הסיכוי שאזורים בעיר חווים עקבות יובש. ניתן ליישם מודל זה על מספר מערכות של תצלומי לווין, כאשר המוצלח מביניהם הוא S-ENVEN (שסיים את פעולתו לאחרונה), אולם ניתן להפעיל אותו גם שירות על נתוני Sentinel-2 שזריםם בפלטפורמות כגון GEE (Google Earth Engine).

פותחה אפליקציה על בסיס streamlit המנגישה את מודל הסיווג *SVM* שפותח במחקר זה. אףן השימוש באפליקציה מפורט **בנספח 1** למטה. את הקישור לאפליקציה ניתן למצוא כאן:
[/https://drought-risk-ml-analyzer.streamlit.app](https://drought-risk-ml-analyzer.streamlit.app)

מסד נתונים

מסד הנתונים מכיל קובץ Excel עבור המדידות בשטח וдинומי הרחפן (274 ערוצים היפרספקטרליים) הנמצאים על שרת מחשבי-העל של האוניברסיטה העברית. ניתן לקבל את כל הנתונים באופן מסודר מהמעבדה של דוד הלמן.

פירוט המשימות בהתאם לתוכנית המחקר

שם המשימה * יש לציין את שם המשימה ואם היא מקורית או חדשה.	התקדמות	מטרות שהוצבו ולא הושגו	סיבה לשינויים במשימה / או עיכובים בביצוע	שנה
צלום מרחפן לאורך העונה בחלקוות הניסוי, עיבוד הנתונים להחזריות והפקת מדדים ספקטרליים	בוצע	מספר חודשיים עם נתונים חלקיים או חסרים	תקולות טכניות של הרחפן	א
איסוף נתונים פמ"ע -AI&LA במינימ השוניים בחלוקת המעורבת	בוצע			
איתור חלקות הומוגניות, חד- מיניות באזור	בוצע			
איסוף נתונים פמ"ע -AI&LA במינימ בחלקוות חד-מיניות	בוצע (לא דוח Cain)			
איסוף נתונים מטאורולוגיים בחלקוות	נתונים זמינים (תחנת בית ג'ימאל)			
העמדת ניסוי הפחחת גשם בחלוקת המורבת	בוצע			א
ניטור על-ידי חיישי תכולת מים	בוצע			א בקירקע

			בוצע	השוואת ממדדי הצומח עם המדדים הספקטראליים מהרחפן בחלוקת ה nisiוני	א
			בוצע	ניתוח גורמים משפיעים (בעזרת הנתונים המטאורולוגיים)	ב
			בוצע	שימוש בימידת מכונה לניתוח ומודול ממדדי צומח ומדדים ספקטראליים	ב
מניתוח הנתונים ותוצאות המידול הסתבר שנתוני הסניטר VEN תוצאה טוביה יוטר	הורדו נתוני VEN	בוצע בשנה ג (עיכוב בשל המלחמה)		הורדת נתונים מלוויינים - Sentinel 2 ו- Planet לאזור החלקות	ב
נתוני AI שנאספו נמצאו ברמה נמוכה מד".	AI לא מודל	בוצע בשנה ג (עיכוב בשל המלחמה)		מודול פמ"ע ו-AI בעזרת נתוני הלוויינים, הרחפן, והמידות שימוש בימידת מכונה לצורך ניתוח גורמים מסבירים על-סמן נתוני החלקה	ב
עיכובים בגלל המלחמה.		בוצע בשנה ג (עיכוב בשל המלחמה)		איסוף נתונים אינונטיר של יערות קק"ל ניתוח ראשוני של ממדדים ספקטראליים ביערות קק"ל איסוף מידע מטאורולוגי מתחנות ובעזרת תוצריו קיימים (CHIRPS, CHIRTS) עבור היערות בישראל	ב
			בוצע	אימות המודל מול נתונים החלקות בשנה השלישית פיתוח המודל ליערות קק"ל לאורך משך הזמן בישראל ניתוח גורמים מקיף	ג
נתוני VEN זמןנים לטוויה מוגבל של שנים.	הושווה מול שנים מקדמות יותר	בוצע		שימוש המודל לחיזוי (מייפוי) אזורי עקה ביערות קק"ל עבור שנים קדומות השוואה מול השנה האחרונה	ג

Fishman, N., Yungstein, Y., Yaakobi, A., Obersteiner, S., Rez, L., Mulero, G., Michael, Y., Klein, T., & Helman, D. (2025). **Leaf Water Potential in a Mixed Mediterranean Forest from Machine Learning and Unmanned Aerial Vehicle (UAV)-Based Hyperspectral Imaging.** *Remote Sensing*, 17(1), 106. <https://doi.org/10.3390/rs17010106>

Yungstein Y, Fishman N, Lerner G, Mulero G, Michael Y, Yaakobi A, Obersteiner S, Rez L, Klein T, Helman David. **Early detection of drought-stressed stands in Mediterranean forests using machine learning classification models and a rainfall exclusion experiment.** (*Under review*).
Preprint: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5292430.

כינויים

דצמבר 2022: פוסטר בכנס **מחקרי יער ושטחים פתוחים**, תל אביב – "השפעת שינוי אקלים על תנאי עיקת יובל ביערות ישראל: הערכה לטוחה הקצר והארוך באמצעות שיטות חישה מרוחק"

אוגוסט 2024: הרצאה בכנס הבינלאומי **ForestSAT 2024**, רוטורואה, ניו זילנד – "Detecting drought stress in key Mediterranean woody species using machine learning and hyperspectral drone imagery"

נספח 1 – הפעלת אפליקציית מודל הסיווג

סקירה כללית

פותחה אפליקציה web אינטראקטיבית המנגישה את המודלים להערכת סיכון בצורת באמצעות תמונות לוויין. האפליקציה, שפותחה בסביבת Streamlit, מאפשרת למשתמשים לישם את המודלים המאומנים על נתוני לוויין-חדרים ללא צורך במידע תכנות או עיבוד נתונים מתקדם. המערכת תומכת בשני סוגי נתונים לוויין - Sentinel ו-Venue - 2 ומספקת משקל פשוט וידידותי להפעלת המודלים ולצפיה בתוצאות.

אופן השימוש באפליקציה

השימוש באפליקציה מתחילה בבחירה סוג הלויין המתאים לנtíנים הקיימים. המשמש בוחר בין לוויין Venue או Sentinel-2 בהתאם לנtíנים שברשותו. לאחר הבחירה, הוא מעלה קובץ GeoTIFF רב-ערוצי המכיל את נתונים הלויין. המערכת בודקת אוטומטית שהקובץ מכיל את מספר העורצים הנדרש ומציג הودעות מתאימות במקרה של בעיה (איור נ1).

כאשר הקובץ נטען בהצלחה, האפליקציה מפעילה את המודל המתאים ומבצעת עיבוד של התמונה. מאחר וכל מודל מייצר תחזית הסתברותית בין 0-1 לסיכון לוויין, המשתמש יכול לקבוע ערך סף שימושי הסיווג הבינארי של המודל יקבע כיובש. זה מאפשר לשולט על רגישות התוצאות ועל רמות הסיכון שמצוגות כפלט. עם התהילה מוצג למשתמש עם סרגל התקדמות, והוא יכול לעקוב אחר מצב העיבוד. המערכת מחלקת את התמונה לקטעים קטנים לעיבוד עיל, ומפרקה עבור כל פיקסל הסתברות לסיכון בצורת.

Satellite-Based Drought Risk Assessment

Citation: If you use this application or the model's outputs in your research, please cite:

Yungstein, Y., Fishman, N., Lerner, G., Mulero, G., Michael, Y., Yaakobi, A., Obersteiner, S., Rez, L., Klein, T., & Helman, D. (2025). Early detection of drought-stressed stands in Mediterranean forests using machine learning classification models and a rainfall exclusion experiment

This application leverages advanced machine learning methods to estimate drought risk from satellite imagery. It integrates trained Support Vector Machine models that interpret multi-band geospatial data to produce pixel-wise probability maps of drought vulnerability.

Key Features:

- **Multi-Satellite Support:** Process imagery from both Venus and Sentinel-2 satellites.
- **High-Quality Visualization:** RGB composites from appropriate satellite bands.
- **Drought Probability Mapping:** Pixel-level probability assessments of drought risk.
- **Robust Statistical Analysis:** Histograms, thresholds, and summary statistics for intuitive interpretation.
- **Overlay Functionality:** Superimpose drought risk areas over RGB images for contextual insights.

Select Satellite Data Source

Choose the satellite data source that matches your input GeoTIFF file:

Select one or more satellite data sources:

Venus  Sentinel-2 



Upload Satellite GeoTIFF

Please upload a multi-band GeoTIFF file from one of your selected satellites. The appropriate model will be applied based on your selection.

Choose a GeoTIFF file



Drag and drop file here

Limit 200MB per file • TIF, TIFF

Browse files

איך נ1. דף הנחיתה של האפליקציה. המשתמש מקבל הנחיות מפורטות לגבי אופן השימוש ושלבי הפעולה.

לאחר סיום העבודה, האפליקציה מציגה את התוצאות באربע תצוגות שונות. התצוגה הראשונה מציגה את התמונה בצבעים טבניים, המאפשרת למשתמש לזהות את האזור הנחקר. התצוגה השנייה מציגה מפת הסטברות צבעונית המדגישה את אזורי הסיכון השונים, כאשר המשתמש יכול לבחור בין מפות צבע שונות ולהתאים את הסוף להגדרת אזורי סיכון גובה. התצוגה השלישית מספקת נגישות סטטיסטית כולל היסטוגרמה של התפלגות ההסתברויות וחישוב השטחים הכלולים. התצוגה הרביעית משלבת את התמונה הטבעית עם הדגשה של אזורי הסיכון הגבוה (איך נ2).

Analysis Results

RGB Image Probability Map Statistical Analysis Overlay

Venus RGB + Forecast Overlay (High-Risk Areas)

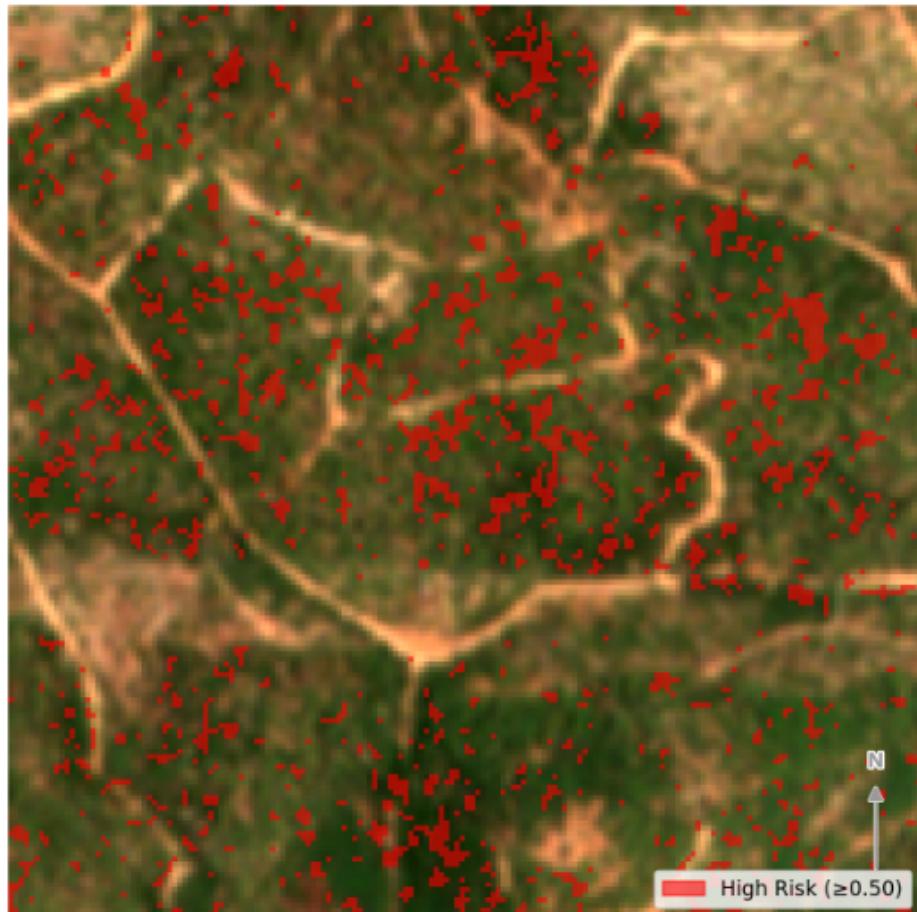
Set Forecast Layer Transparency

0.60

0.00

1.00

Venus RGB + Drought Risk Overlay



איור נ.2. דוגמה לפלט של המודל כפי שIALIZED בציגות בסיס RGB של הדימות שהמשתמש העלה יחד עם הדגשה של הפיקסלים שבו רום המודל חזה סיכון ליבש על בסיס ערך הסף שהמשתמש בחר. בדוגמה זאת ניתן לראות את הסיכון בעיר ישי, על בסיס תמונה לוון Venus מתאריך 18 בספטמבר 2019 עם ערך סף של 0.5.

בסיום התהיליך, המשתמש יכול להוריד את התוצאות בפורמט CSV המכיל את כל ערכי ההסתברות, והשני הוא קובץ GeoTIFF המשמר את המידע הגיאוגרפי ומאפשר המשך עבודה במערכות GIS. האפליקציה מספקת גם הנחיות שימוש מפורטות. האפליקציה מיועדת לחוקרם ולמומחים הэкולוגים לכלי פשוט ונגישלי לשימוש המודלים על אזורים חדשים. היא מאפשרת שימוש מעשי במודלים ללא צורך במידע טכני בתכונות או בעיבוד תמונה, ומספקת תוצאות מיידיות וברורות שניתן להשתמש בהן להחלטות מקצועיות בתחום ניהול יערות ומשאבי טבע.

גופח 2 – מאמר 1 (Remote Sensing)

Fishman, N., Yungstein, Y., Yaakobi, A., Obersteiner, S., Rez, L., Mulero, G., Michael, Y., Klein, T., & Helman, D. (2025). **Leaf Water Potential in a Mixed Mediterranean Forest from Machine Learning and Unmanned Aerial Vehicle (UAV)-Based Hyperspectral Imaging.** *Remote Sensing*, 17(1), 106. <https://doi.org/10.3390/rs17010106>

Article

Leaf Water Potential in a Mixed Mediterranean Forest from Machine Learning and Unmanned Aerial Vehicle (UAV)-Based Hyperspectral Imaging

Netanel Fishman ¹, Yehuda Yungstein ^{1,2}, Assaf Yaakobi ³, Sophie Obersteiner ³, Laura Rez ³, Gabriel Mulero ^{1,2}, Yaron Michael ¹, Tamir Klein ³ and David Helman ^{1,2,*}

¹ Department of Soil and Water Sciences, Institute of Environmental Sciences, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 7610001, Israel; netanel.fishman@mail.huji.ac.il (N.F.); yehuda.yungstein@mail.huji.ac.il (Y.Y.); gabriel.mulero@mail.huji.ac.il (G.M.); yaron.michael@mail.huji.ac.il (Y.M.)

² The Advanced School for Environmental Studies, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 9190401, Israel

³ Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel; assaf.yaakobi@weizmann.ac.il (A.Y.); sophie.obersteiner@weizmann.ac.il (S.O.); laura.rez@weizmann.ac.il (L.R.); tamir.klein@weizmann.ac.il (T.K.)

* Correspondence: david.helman@mail.huji.ac.il

Abstract: Leaf water potential (ψ_{leaf}) is a key indicator of plant water status, but its measurement is labor-intensive and limited in spatial coverage. While remote sensing has emerged as a useful tool for estimating vegetation water status, ψ_{leaf} remains unexplored, particularly in mixed forests. Here, we use spectral indices derived from unmanned aerial vehicle-based hyperspectral imaging and machine learning algorithms to assess ψ_{leaf} in a mixed, multi-species Mediterranean forest comprised of five key woody species: *Pinus halepensis*, *Quercus calliprinos*, *Cupressus sempervirens*, *Ceratonia siliqua*, and *Pistacia lentiscus*. Hyperspectral images (400–1000 nm) were acquired monthly over one year, concurrent with ψ_{leaf} measurements in each species. Twelve spectral indices and thousands of normalized difference spectral index (NDSI) combinations were evaluated. Three machine learning algorithms—random forest (RF), extreme gradient boosting (XGBoost), and support vector machine (SVM)—were used to model ψ_{leaf} . We compared the machine learning model results with linear models based on spectral indices and the NDSI. SVM, using species information as a feature, performed the best with a relatively good ψ_{leaf} assessment ($R^2 = 0.53$; RMSE = 0.67 MPa; rRMSE = 28%), especially considering the small seasonal variance in ψ_{leaf} ($\pm\sigma = 0.8$ MPa). Predictions were best for *Cupressus sempervirens* ($R^2 = 0.80$) and *Pistacia lentiscus* ($R^2 = 0.49$), which had the largest ψ_{leaf} variances ($\pm\sigma > 1$ MPa). Aggregating data at the plot scale in a ‘general’ model markedly improved the ψ_{leaf} model ($R^2 = 0.79$, RMSE = 0.31 MPa; rRMSE = 13%), providing a promising tool for monitoring mixed forest ψ_{leaf} . The fact that a non-species-specific, ‘general’ model could predict ψ_{leaf} implies that such a model can also be used with coarser resolution satellite data. Our study demonstrates the potential of combining hyperspectral imagery with machine learning for non-invasive ψ_{leaf} estimation in mixed forests while highlighting challenges in capturing interspecies variability.



Academic Editor: Junjun Jiang

Received: 21 November 2024

Revised: 22 December 2024

Accepted: 25 December 2024

Published: 31 December 2024

Citation: Fishman, N.; Yungstein, Y.; Yaakobi, A.; Obersteiner, S.; Rez, L.; Mulero, G.; Michael, Y.; Klein, T.; Helman, D. Leaf Water Potential in a Mixed Mediterranean Forest from Machine Learning and Unmanned Aerial Vehicle (UAV)-Based Hyperspectral Imaging. *Remote Sens.* **2025**, *17*, 106. <https://doi.org/10.3390/rs17010106>

Copyright: © 2024 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

Keywords: NDVI; random forest; remote sensing; SVM; water; XGBoost

1. Introduction

Forests increasingly face harsh conditions, leading to worldwide tree mortality in the last few decades [1]. This trend is expected to aggravate under projected climate change. Rising temperatures and altered precipitation patterns will induce severe drought conditions in many regions [2]. Despite their adaptation capacity, many forest species might not cope with such conditions, leading to a decline in growth rates and increased mortality [3,4]. Monitoring forests' water status is vital to facing such expanded adverse effects. This requires developing accurate, efficient, and large-scale assessment methods [5].

One of the most valuable metrics for characterizing drought stress and water status in woody species is the leaf water potential (ψ_{leaf}). ψ_{leaf} represents the degree of stability of the water in the xylem transport conduits of the trees by becoming more negative as the drought stress increases [6]. ψ_{leaf} is a key indicator of plant water status and can provide valuable insights into the severity of drought stress experienced by the tree [7]. However, measuring ψ_{leaf} in the field is a time-consuming and labor-intensive process, typically involving the use of pressure chambers on individual leaves [8]. This approach limits the spatial coverage and temporal resolution of ψ_{leaf} measurements, making it challenging to capture the heterogeneity of drought stress within and across forest stands.

Remote sensing offers a promising solution to this challenge, enabling the rapid and non-invasive assessment of vegetation status across large areas [9,10]. Hyperspectral reflectance information obtained from satellites or drone-based sensors has been widely used to estimate various plant traits, including the leaf area index, chlorophyll content, water content, transpiration rate, and water use efficiency [11–19]. The physical basis for using spectral information to estimate plant traits is based on the interaction of light with plants' biophysical and biochemical characteristics. Plant traits corresponding to water content or stress can directly influence absorption in the NIR region, particularly around the 970 nm and 1200 nm water absorption bands, as well as in the SWIR, particularly 1450, 1900, and 1950 nm [20]. Other indirect responses of water-related stress that may lead to variations in pigment concentrations within the plant's intercellular organelles primarily affect reflectance in the visible range [21].

Thus, several studies have explored the potential for using spectral vegetation indices, a mathematical combination of two or more bands derived from remote sensing data, to assess ψ_{leaf} across vegetation types. Raj et al. [22] detailed several water absorption bands in the visible and NIR regions useful for monitoring water-related variables like ψ_{leaf} . They identified seven bands in the 400–1000 nm, which are sensitive to vibrational overtone frequencies of O-H bonds of water molecules, and created four normalized indices that showed a high correlation to the leaf water content of Maize crops. Peñuelas et al. [23] developed the water band index based on near-infrared (NIR) reflectance (R_{970}/R_{900}) to estimate ψ_{leaf} under salinity stress conditions. Stimson et al. [24] demonstrated that drought stress in two tree species (*Pinus edulis* and *Juniperus monosperma*) could be quantified using indices incorporating 980 nm and 1200 nm wavelength bands, significantly correlated with the measured ψ_{leaf} . Italiano et al. [25] showed corresponding temporal patterns that had a positive linear relationship with earlywood hydraulic diameter (linked to water conductivity) in drought-prone Mediterranean forests using common vegetation indices such as the NDVI, EVI, and NDWI.

Additionally, Othman et al. [26], analyzing Landsat satellite imagery, found that the shortwave-infrared (SWIR) band ratio effectively predicted ψ_{leaf} in pecan orchards. More recently, using three SWIR bands, Wang et al. [27] proposed a modified vegetation index ($(R_{1740}-R_{2370})/(R_{1740}-R_{1750})$), which showed a strong correlation with ψ_{leaf} across different leaf structures. Such a myriad of studies, among others, suggest that remotely sensed spectral data could be successfully used for monitoring ψ_{leaf} in complex systems such as

mixed forests. Yet, such complex systems may require models that account for nonlinear relationships between the remotely sensed data and ψ_{leaf} .

Machine learning techniques have shown great promise for modeling plant water content and status using remote sensing data. Machine learning algorithms can effectively capture the complex and nonlinear relationships between spectral reflectance patterns and plant physiological traits (e.g., Li et al. [28]). Several studies have successfully applied machine learning algorithms such as random forests (RFs), support vector machine (SVM), and artificial neural networks (ANNs) to predict leaf water content, relative water content, and equivalent water thickness from hyperspectral and multispectral imagery [29–31].

Despite the progress made in applying remote sensing to assess ψ_{leaf} , there remains a gap in our understanding of how these techniques perform in diverse, mixed-species forests at high spatial resolutions (<1 m). To the best of our knowledge, no study has combined high-spatial-resolution remote sensing data acquired from sensors onboard unmanned aerial vehicles (UAVs) with machine learning algorithms to predict ψ_{leaf} in mixed forest stands. Many Mediterranean forests, for example, are characterized by a complex mosaic of woody species with varying drought tolerances and water use strategies [32,33]. To accurately assess the water status of these forests, it is necessary to develop remote sensing methods that can account for such heterogeneity and species-specific ψ_{leaf} at fine spatial scales [34]. Then, generated ψ_{leaf} maps could be used to point at specific stands or stand areas under stress conditions, allowing for precursory intervention. This is particularly important in places like Israel, where about half of the forested areas are planted and managed by the Jewish National Fund (JNF [35]). Such monitoring can help detect stands under stress and treat the stands with irrigation or thinning the stand density under pressure [36,37].

Here, we address the above research gap by examining the use of high-spatial-resolution UAV-derived hyperspectral imagery taken over a year in open-field experimental plots to assess ψ_{leaf} in five key Mediterranean woody species. We developed machine learning models based on the hyperspectral data and field ψ_{leaf} measurements and compared the models with commonly used spectral indices to assess ψ_{leaf} . We evaluated the ability of spectral reflectance data and machine learning models to capture plot and species-specific ψ_{leaf} changes. We hypothesize that (i) ψ_{leaf} would be easier to model for species with a broader range of ψ_{leaf} values than species displaying a narrow range of values. We also hypothesize that (ii) machine learning algorithms should significantly enhance the ability to predict ψ_{leaf} compared to simple linear regression models using spectral indices, particularly those based on nonlinear relationships.

2. Data and Methods

2.1. Study Site and Experimental Design

The experiment was conducted in Yishi Forest, a semi-arid mixed Mediterranean forest in the Judean foothills, Israel [34,38]. The forest is located 4 km southwest of Beit Shemesh, Israel ($31^{\circ}43'N$ $34^{\circ}57'E$, Figure 1). It covers an area of ~650 ha with an average elevation of 300 m a.s.l. and a mean annual precipitation of 460 mm yr^{-1} (annual mean of the last 20 years). Precipitation occurs mainly from November to May. The mean annual temperature is $20.4 \pm 6.8\text{ }^{\circ}\text{C}$, with mean winter (January–March) and summer (June–August) temperatures of $16.5 \pm 1.9\text{ }^{\circ}\text{C}$ and $24.1 \pm 7.9\text{ }^{\circ}\text{C}$, respectively [39]. The predominant soil type in Yishi is terra rossa, which consists of A and C soil horizons, with C horizon soil penetrating the cracks between the weathered limestone bedrock. Soil depth (A horizon) is ~21 cm, ranging from 16 to 25 cm [33].

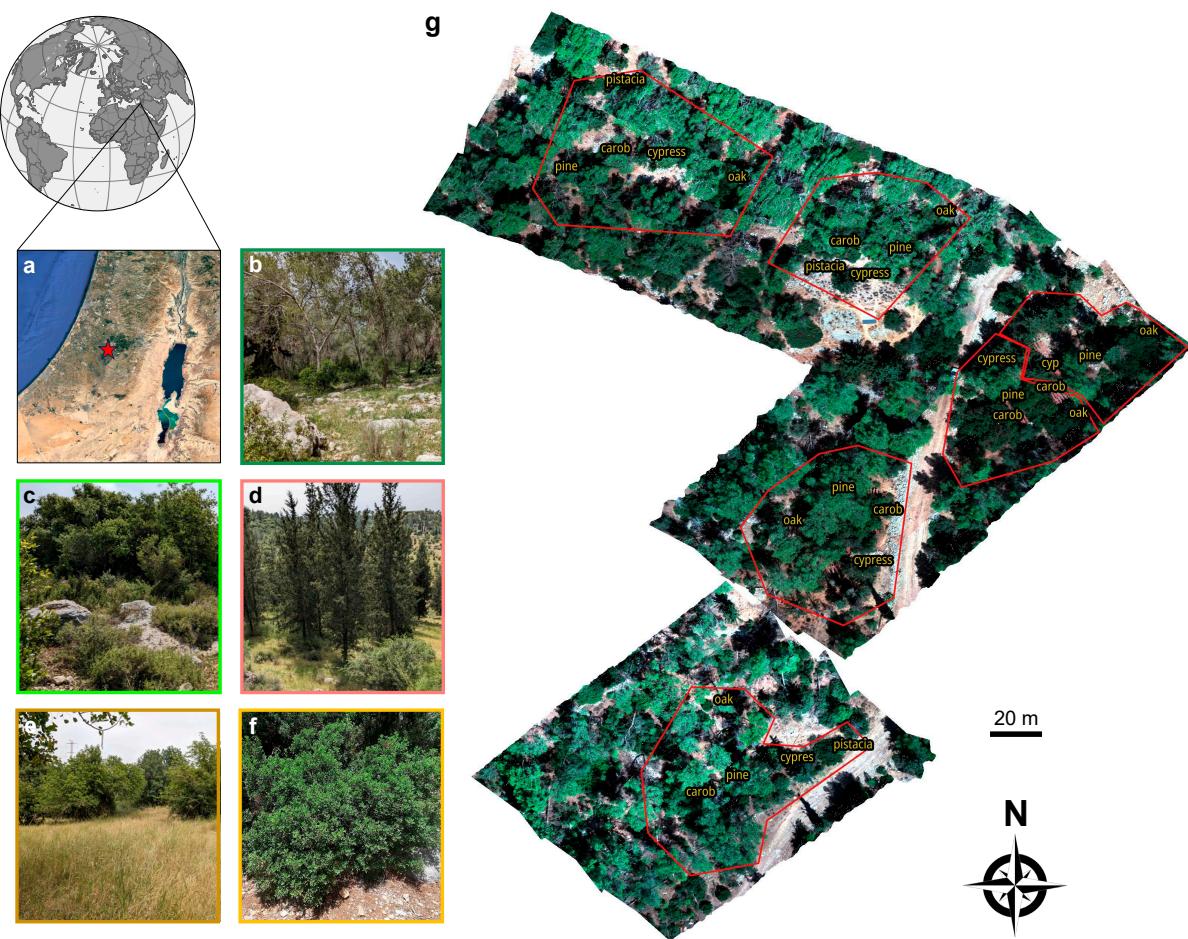


Figure 1. (a) The study area (marked by a red star) and the five key Mediterranean woody species comprising Yishi Forest, which includes (b) Pine (*Pinus halepensis*), (c) Oak (*Quercus calliprinos*), (d) Cypress (*Cupressus sempervirens*), (e) Carob (*Ceratonia siliqua*), and (f) Pistacia (*Pistacia lentiscus*). (g) An RGB image derived from the hyperspectral camera onboard an M600 Pro UAV showing the six plots in the studied area.

The vegetation in Yishi is dominated by the planted gymnosperm woody species *Pinus halepensis* (Figure 1b) and *Cupressus sempervirens* (Figure 1d), as well as the local Mediterranean angiosperm woody species *Quercus calliprinos* (Figure 1c), *Ceratonia siliqua* (Figure 1e), and *Pistacia lentiscus* (Figure 1f). In what follows, we will refer to these key woody species by their common English names: Pine, Cypress, Oak, Carob, and Pistacia. These woody species were formerly studied in situ regarding their water relations and carbon management ([33,38], respectively). The forest understory supports a variety of annual plants that thrive from winter to spring.

Six plots, each ~0.05 ha in area, were established within the forest, comprising the five co-occurring woody species (Figure 1g). The plots were divided into two treatments, three plots subjected to rainfall reduction to simulate drought conditions and three control plots under normal rainfall conditions. The rainfall reduction treatment involved installing an open-pipe harvesting system to divert approximately 50% of incident precipitation from the treated plots. Nevertheless, considering the drought-adapted woody species studied here, treatment effects were not expected within the first few years, as reported for a nearby site [40]. Therefore, treatments were not considered in the analysis of this study, but rather, data from both treatments were used together to broaden the range in ψ_{leaf} to enable more robust modeling.

2.2. Field Measurements

ψ_{leaf} was measured using the pressure chamber technique [41]. This method involves enclosing a leaf petiole in a sealed chamber and gradually increasing the pressure inside the chamber until the sap is observed emerging from the cut surface of the petiole. The pressure at which this occurs is equal to the negative of the leaf water potential, providing a direct measurement of the water status within the leaf [42]. Measurements were conducted monthly between October 2021 and March 2023 from 11:00 AM to 1:00 PM to capture maximum daily ψ_{leaf} values across the season. ψ_{leaf} was measured once per date in all the 30 trees used in this study (five species in six plots). In a preliminary trial, we measured five leaves per individual tree on a subset of the studied trees. Variations in ψ_{leaf} among leaves sampled from the same tree were in the magnitude of 2–7%. A major consideration of the field measurements was to complete all measurements within the shortest time possible to avoid significant diurnal variations among the samples. This is paramount when comparing individuals and species in a single day. Thus, to complete the measurements of all 30 trees within 60–90 min (representing similar temperature, humidity, and radiation conditions), having a single pressure chamber instrument and two workers, and following the small ψ_{leaf} variability observed in our trial, we opted to sample a single leaf per tree (per date). Excised leaves were immediately placed in airtight plastic bags and kept cool to minimize measurement errors from time lags. A total of 30 leaves were sampled per date and measured using a PMS1515 pressure chamber (PMS, Albany, OR, USA). The total number of sampled leaves during the study period was 480 (16 months, 30 leaves per date).

2.3. UAV Platform and Spectral Data Acquisition

Hyperspectral imagery was acquired monthly concurrent with the ψ_{leaf} measurements starting from March 2022 to February 2023, using a Nano-Hyperspec camera (Headwall Photonics, Boston, MA, USA) mounted on a DJI Matrice 600 Pro (M600) Hexacopter. The NanoSpec sensor is a push-broom hyperspectral nano-sensor with 274 spectral bands and 640 spatial pixels within the visible-near-infrared range from 400 nm to 1000 nm. The M600 is controlled via a handheld remote control transmitter and a ground control station with a navigation data link, which sends waypoint navigation information to the aircraft from a laptop computer. The M600 carries a Global Navigation Satellite System (GNSS) and an Inertial Measurement Unit (IMU) (SBG Systems North America, Inc., Chicago, IL, USA).

The images were collected between 11:00 AM and 1:00 PM, concurrent with the ψ_{leaf} measurements, at the height of 60 m above ground level, providing a spatial resolution of 2–3 cm per pixel. Three flights were required to cover the study area. For radiometric calibration, a 3 m by 3 m in situ gray-white reflectance calibration with three strips and distinct reflectance factors (56%, 30%, and 11% reflectance) was set up within the flight scene in each campaign (date). For geometric corrections, a ground-based GNSS receiver—Trimble SPS585 precision RTK (Trimble Inc., Westminster, CO, USA)—was used to collect static geolocation data to calculate the post-processing kinematic (PPK) flight trajectory using a smoothed best-estimated trajectory (SBET) file generated from the POSpac UAV™ Version 8.9 software tool (Applanix, Richmond Hill, ON, Canada).

2.4. Spectral Data Processing

Radiometric calibration, geometric corrections, and ortho-mosaicing were applied to the raw hyperspectral image cubes via the SpectralView software version 3.1.4 (Headwall Photonics, Fitchburg, MA, USA). Shades and non-vegetated areas within the image (Figure 2a) were masked using the Quantum Geographic Information System (QGIS version 3.32.3, Free Software Foundation, Boston, MA, USA). First, the normalized difference vegetation index (NDVI) was calculated, and pixels with NDVI < 0.3 were excluded to

eliminate soil and understory vegetation (Figure 2b). Then, we used near-infrared (NIR) reflectance values below a threshold of 0.07–0.2 to eliminate shaded parts in the canopy since these affect the spectral signal despite changes in ψ_{leaf} . We found these thresholds suitable for that purpose following trial and error after visually inspecting the clipped RGB (Figure 2c).

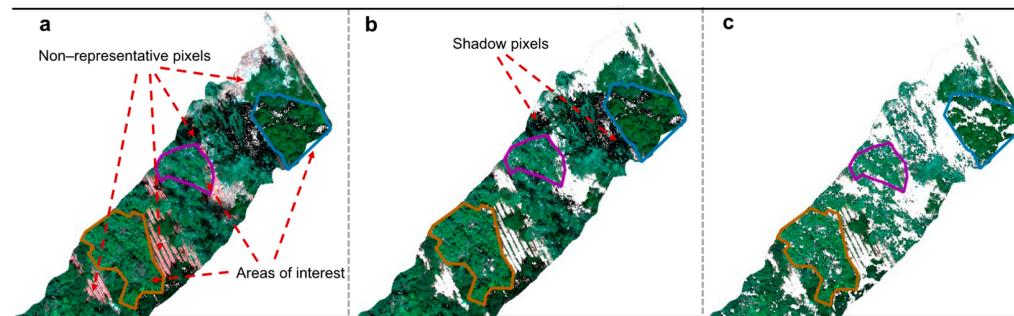


Figure 2. An example of masking non-representative pixels within the pre-processing pipeline. (a) The original image with polygons of the areas of interest (i.e., the targeted trees) and non-vegetated pixels within the image. (b) The image after masking non-representative pixels using an NDVI threshold of <0.3 and (c) shaded pixels with an NIR reflectance threshold.

The reflectance values of retained pixels were averaged for each tree canopy, which was manually delineated, to obtain a single mean reflectance spectrum representing the entire canopy. The Savitzky–Golay filter [43], with a window size of 20 bands and second-order polynomials, was applied to eliminate potential artifacts and anomalies in the spectral signatures caused by measurement errors (Figure 3).

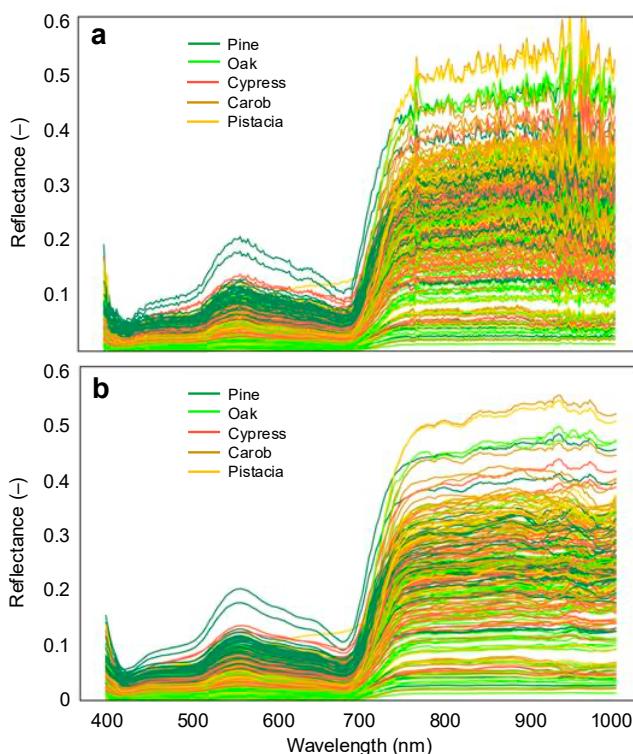


Figure 3. Spectral signatures in the 400–1000 nm range (a) before and (b) after applying the Savitzky–Golay filter smoothing. Each line in the graph corresponds to the averaged signature over all pixels in the canopy per each species and date.

After the spectral data processing, the final dataset used within the modeling pipeline resulted in a sample size of $N = 246$. The data were randomly split into a 70% training set and a 30% test set for model training and evaluation purposes. This ratio is considered useable and generic within an ML pipeline, especially if one wants a larger data sample to evaluate the model and ensure a more robust transferability to other locations and cases.

3. Predictive Models

We followed two approaches to predict ψ_{leaves} from the hyperspectral data. First, we used a simple regression approach, whereas the predictor is a spectral index, whether derived in previous studies or generated in this study for all species together as well as for individual species (see Section 3.1). For the same datasets, we also tested the best NDSI, adding one index at a time within a multivariable linear regression (MLR) model. Second, we used the best NDSIs generated from the hyperspectral data in three machine learning algorithms (Section 3.2).

3.1. Spectral Indices

Twelve known spectral indices within the 400–1000 nm range were tested by conducting a simple linear regression with ψ_{leaf} . The selected indices were based on their ability to directly and indirectly predict water-related plant traits [44–47]. The normalized difference vegetation index (NDVI), the photochemical reflectance index (PRI), and the enhanced vegetation index version two (EVI2) were tested for their ample use in monitoring vegetation status and dynamics, which showed to be indirectly related to water stress through biophysical changes [44,48,49]. COSBNDI, FOSBNDI, SAPSBNDI, and WASCOSBNDI, which were developed based on seven identified bands sensitive to the vibrational overtone frequencies of O-H bonds of water molecules, were tested due to their strong link to leaf water content [22]. We also used the normalized difference water stress index (NDWSI) and the normalized difference water index (NDWI) because they have been shown to be sensitive to plant water status via canopy structural changes [44,50]. There are several versions of the NDWI [19,51,52]. Here, we used the one based on the 540 and 803 nm bands (Table 1), which incorporates the responses from the visible region of the spectrum (the green band) with canopy structural changes within the NIR region to water-related responses since our hyperspectral data were restricted to the VIS–NIR wavelength range (400–1000 nm). Finally, the water index (WI), developed specifically to assess ψ_{leaf} , was tested. The WI gives information about canopy water status with a change in the 970 nm relative to the 900 nm as the water status changes in the leaves [23]. Table 1 summarizes the 12 spectral indices, providing their mathematical formulation.

Table 1. Previously published spectral indices used in this study to assess ψ_{leaf} .

Index	Name	Formula	Reference
NDVI	Normalized difference vegetation index	$\frac{\rho_{\text{NIR}} - \rho_R}{\rho_{\text{NIR}} + \rho_R}$	[53]
PRI	Photochemical reflectance index	$\frac{\rho_{531} - \rho_{570}}{\rho_{531} + \rho_{570}}$	[54]
940/960	Reflectance ratio of 940 and 960 nm	$\frac{\rho_{940}}{\rho_{960}}$	[55]
940/960/NDVI	Reflectance ratio of 940/960 nm and NDVI	$\frac{\rho_{940} / \rho_{960}}{\text{NDVI}}$	[55]
EVI2	Enhanced vegetation index 2nd version	$2.5 \left(\frac{\rho_{\text{NIR}} - \rho_R}{\rho_{\text{NIR}} + (6 \rho_R) - (7.5 \rho_B) + 1} \right)$	[56]

Table 1. Cont.

Index	Name	Formula	Reference
COSBNDI	Combined overtone of stretching bands—normalized difference index	$\frac{\rho_{660} - \rho_{420}}{\rho_{660} + \rho_{420}}$	[22]
FOSBNDI	Forth overtone of stretching bands—normalized difference index	$\frac{\rho_{529} - \rho_{698}}{\rho_{529} + \rho_{698}}$	[22]
SAPSBNDI	Small absorption peak of stretching bands—normalized difference index	$\frac{\rho_{750} - \rho_{970}}{\rho_{750} + \rho_{970}}$	[22]
WASCOSBNDI	Water absorption shoulder due to the combined overtone of stretching bands—normalized difference index	$\frac{\rho_{800} - \rho_{847}}{\rho_{800} + \rho_{847}}$	[22]
NDWSI	Normalized different water stress index	$\frac{\rho_{850} - \rho_{970}}{\rho_{850} + \rho_{970}}$	[50]
NDWI	Normalized different water index	$\frac{\rho_G - \rho_{\text{NIR}}}{\rho_G + \rho_{\text{NIR}}}$	[44]
WI	Water index	$\frac{\rho_{970}}{\rho_{900}}$	[45]

Note: R is red (670 nm), G is green (540 nm), B is blue (480 nm), and NIR is near-infrared (803 nm).

We further leveraged the multiband option of our hyperspectral data (274 bands) and calculated thousands of combinations of a normalized difference index to identify new and possible band combinations suitable for ψ_{leaf} other than the already established 12 spectral indices used in this study. The normalized difference spectral index (NDSI) approach uses a two-band combination of reflectance [57]:

$$\text{NDSI} \text{ (unitless)} = \frac{b1 - b2}{b1 + b2} \quad (1)$$

where b1 and b2 are the reflectance values of any two wavelength bands. Such a normalization mitigates the influence of varying illumination conditions, isolating the relative difference between the selected bands [57].

Since each hyperspectral image contains 274 bands from 400 to 1000 nm, a total of 37,401 NDSI options could be derived from combining each pair of bands (i.e., 274×273 options divided by 2 to avoid duplications with inverted order) per image per individual. Each NDSI value was used in a simple linear regression against the ψ_{leaf} . After this, we selected the top-performing indices yielding the highest coefficient of determination (R^2). First, we excluded indices that showed a high correlation for any two bands close to each other within a 10 nm interval. Then, we chose the index with the best correlation and added four more local maxima indices for all species datasets (Supplementary Figure S1). This process was conducted at the individual tree level, with spectral data analyzed separately for each tree.

A multivariable linear regression (MLR) model was tested with the top five indices, starting with simple regression and adding one index at a time to verify the effect of increasing the number of unique bands on the linear regression model's accuracy.

Finally, we averaged the spectral and ψ_{leaf} values for each plot (comprising five species) for each sampling date. We calculated a new NDSI using these plot-level averages and performed linear regressions between the spectral and the ψ_{leaf} data. This averaging aimed to develop models at the plot scale ('general' model). However, due to the reduction in sample size (fivefold fewer samples), machine learning models could not be applied to this dataset.

3.2. Machine Learning Algorithms

Only the five selected NDSIs exhibiting the highest correlation with ψ_{leaf} were used in three machine learning algorithms: random forest (RF), extreme gradient boosting (XGB), and support vector machine (SVM). We did not include the known spectral indices to avoid model overfitting because some of the known indices had overlapping bands and similar formulations as the NDSIs.

RF is an ensemble learning method that constructs multiple decision trees and aggregates their predictions, accounting for variations and reducing overfitting [58]. XGB is a scalable tree-boosting system that builds an ensemble of weak prediction models sequentially, with each successive model aiming to correct the errors of its predecessor [59]. SVM is a supervised learning algorithm that constructs hyperplanes in high-dimensional space to maximize the margin between classes. It can handle linear and nonlinear relationships through kernel functions, unlike RF and XGB, which are ensemble methods, thus requiring a larger sample size than SVM, which can learn from small samples [60,61]. The key parameters driving each ML algorithm are listed in Supplementary Table S1.

All models were trained on 70% of the data (training set) using fine-tuned hyperparameter (Supplementary Table S1) and cross-validation. The hyperparameter optimization for each model was performed using the scikit-learn randomgridsearchCV module. This module allows for the iterative search for the best combination of the algorithm-specific parameter(s), selecting the best estimators based on the highest-ranking performance of the models fitted. Finally, model performance was evaluated on the remaining unseen 30% (test dataset) using metrics such as the coefficient of determination (R^2), the root mean squared error (RMSE), the ratio of performance to deviation (RPD), and the mean absolute error (MAE). An ensemble model (AVG) was also constructed by averaging the predictions from all individual models. Feature importance analysis was conducted to identify each model's most influential spectral indices.

In a subsequent step, species information was incorporated as a categorical variable using one-hot encoding and added to the input features [62]. The modeling process was repeated, including hyperparameter tuning and performance evaluation on the test set. The RMSE and R^2 were calculated separately for each species to assess model performance across species.

The performance of the three machine learning algorithms (RF, XGBoost, and SVM) was compared with the performance of the linear spectral-based models. In addition, the machine learning algorithms were compared with each other, and their robustness was assessed using the residual prediction deviation (RPD [63,64]). RPD is defined as the ratio of the standard deviation of the actual measured ψ_{leaf} to the RMSE. We adopted the criteria of Mouazen et al. [65] for classifying RPD scores as follows: an RPD below 1.5 indicates that the model is not usable, an RPD between 1.5 and 2.0 indicates a possibility to distinguish between high- and low-value groups, and an RPD between 2.0 and 2.5 makes approximate quantitative predictions possible. The models were classified as good and excellent for RPD between 2.5 and 3.0 or above 3.0, respectively. Figure 4 presents the conceptual modeling framework and schemes.

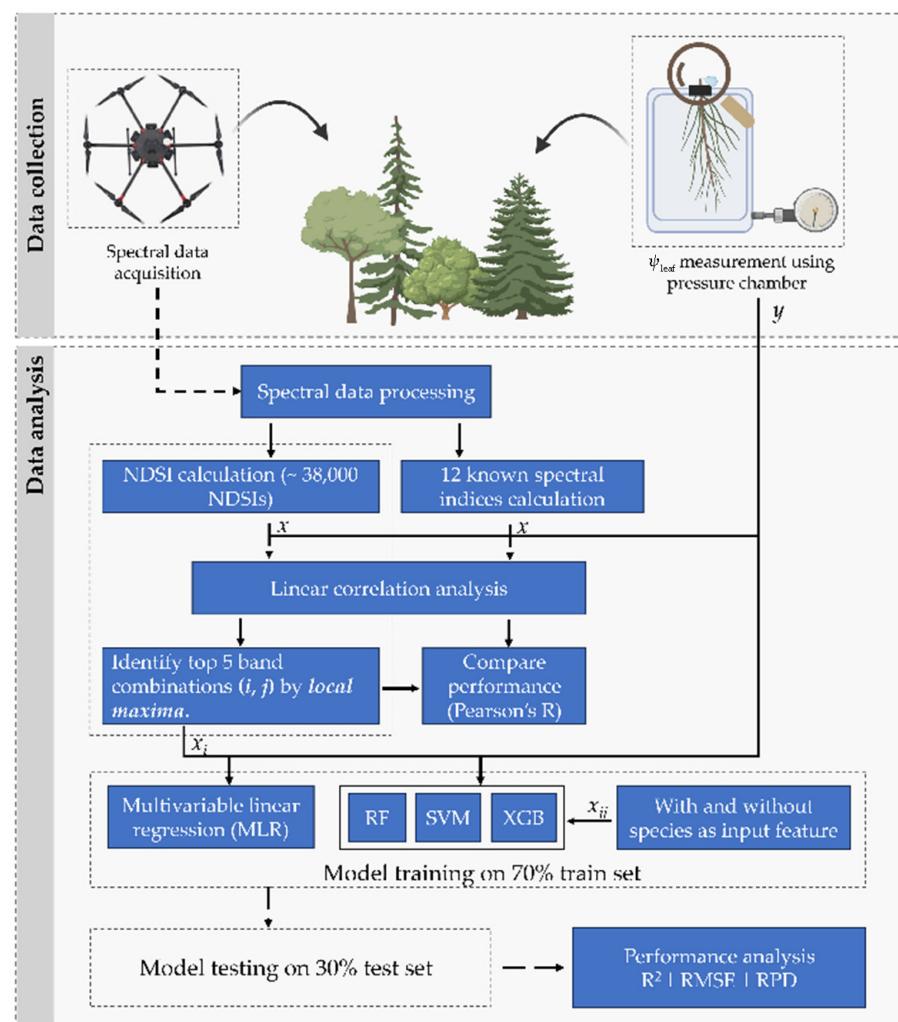


Figure 4. Schematics of the research framework, from the data collection to the machine learning modeling. x and y in the graph mean predictor and predicted variables, respectively. x_i and x_{ii} are the predictors using the top 5 NDSI without and with species as input features, respectively. RF, SVM, and XGB stand for random forest, support vector machine, and extreme gradient boosting algorithms.

3.3. Statistical Analysis

The ground-truth ψ_{leaf} measurements were tested for normality at $p > 0.05$ sample with the Shapiro–Wilk test using the JMP 17 Pro statistical software (SAS Institute). All tests for statistical significance of model performance were performed within the Python pipeline at $p < 0.05$.

4. Results

4.1. Leaf Water Potential Dynamics and Correlation with Spectral Indices

We observed that ψ_{leaf} in the five species fluctuated between -1 MPa and -5 MPa and differed in magnitude as well as in their dynamics along the year (Figure 5). For example, Pine (*Pinus halepensis*) had overall the lowest ψ_{leaf} , with an average value of -2.56 MPa over the year. Pine was also the species with the lowest inter-annual variance in ψ_{leaf} , along with the Carob (*Ceratonia siliqua*). In contrast, Cypress (*Cupressus sempervirens*) and Pistacia (*Pistacia sempervirens*) had both the largest ψ_{leaf} variance over the year, reaching during the dry months a ψ_{leaf} as low as -4.5 MPa.

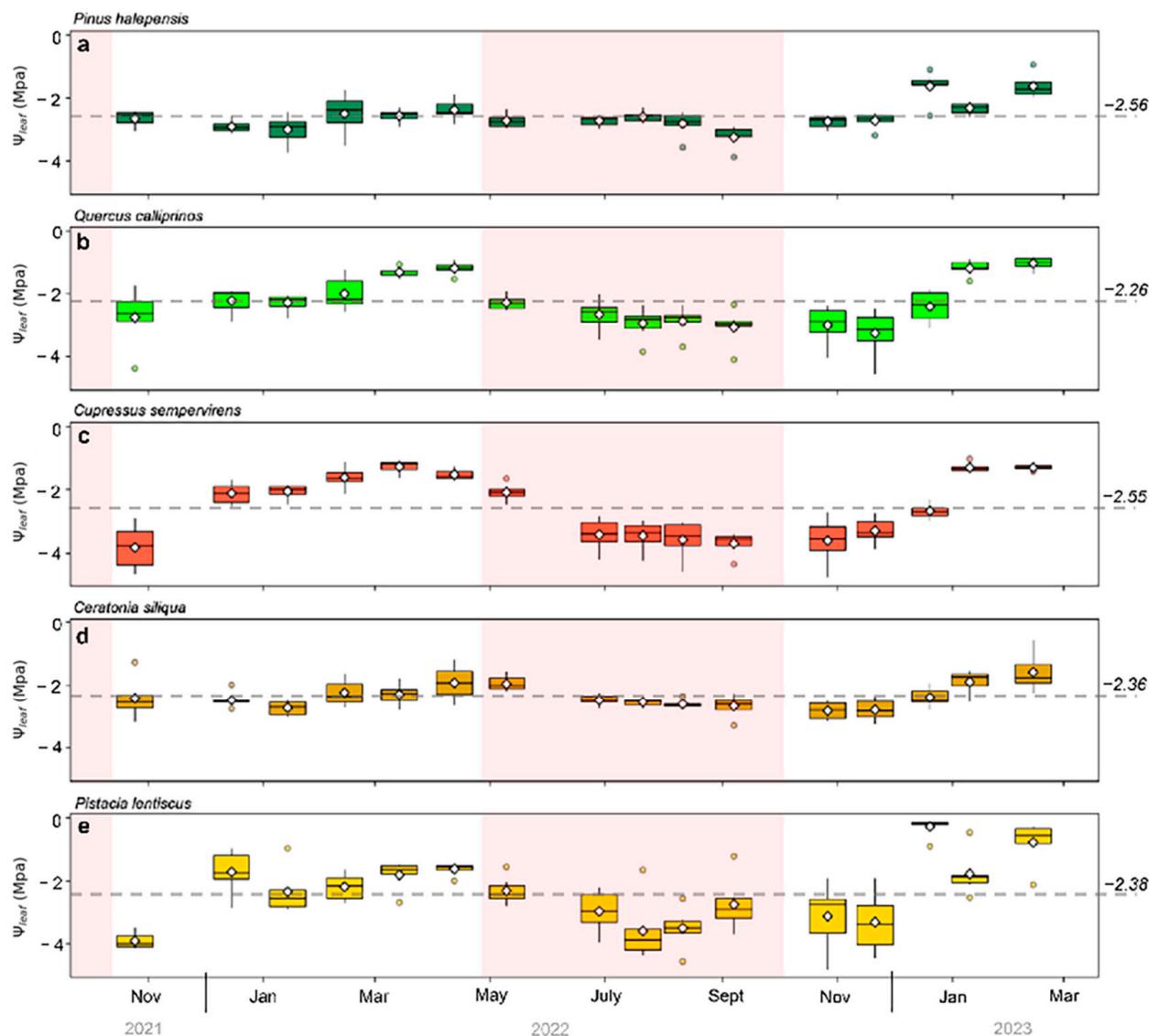


Figure 5. Seasonal variation in ψ_{leaf} throughout the study period across the five key woody species: (a) Pine, (b) Oak, (c) Cypress, (d) Carob, and (e) Pistacia. Each boxplot represents the interquartile range (IQR), with the horizontal line within each box indicating the median and the white diamond symbol the mean. Whiskers extend to the lowest and highest ψ_{leaf} within 1.5 times the IQR, and outliers are displayed as individual points. The horizontal gray dashed line represents the mean value throughout the study period, with its value in MPa next to the line. The pink shaded strips represent dry periods within the year.

Of the twelve known spectral indices, the best linear relationship was achieved with NDVI for Cypress (Pearson's $R = 0.86$), while the other indices displayed only moderate-to-weak correlations with ψ_{leaf} . Pearson's correlations show that some indices had positive and some had negative linear relationships with ψ_{leaf} (Table 2). The correlations did not improve when the best NDSI was used against ψ_{leaf} (with bands 750 nm and 680 nm; Table 2). When considering all species, the linear models with the spectral indices had poor performance, with the best index being the NDVI, together with the best NDSI, both showing a Pearson's R of 0.57.

Table 2. Pearson's correlation coefficient (r) between ψ_{leaf} and spectral indices, including the NDSI, per species, and for all species (All). Highlighted in bold are the highest R^2 scores per index and the second-highest R^2 scores in italics.

Index	Pine	Oak	Cypress	Carob	Pistacia	All
NDVI	0.56	0.70	0.86	0.53	0.54	0.57
PRI	0.31	-0.18	0.19	-0.18	0.55	0.09
940/960	0.20	0.24	0.31	0.29	0.48	0.29
940/960/NDVI	-0.51	-0.65	-0.84	-0.38	-0.34	-0.48
EVI2	0.23	0.10	-0.08	-0.06	0.24	0.08
COSBNDI	-0.32	-0.13	-0.39	-0.18	-0.63	-0.29
FOSBNDI	0.41	0.38	0.63	0.21	0.50	0.43
SAPSBNDI	0.11	0.14	0.49	0.26	0.60	0.27
WASCOSBNDI	-0.22	-0.16	0.26	0.03	0.13	0.03
NDWSI	0.15	0.21	0.50	0.27	0.60	0.29
NDWI	-0.52	-0.65	-0.83	-0.51	-0.67	-0.48
WI	-0.21	-0.27	-0.48	-0.26	-0.60	-0.31
NDSI (680/750)	0.55	0.69	0.86	0.54	0.60	0.57

The MLR model performed slightly better than the spectral indices, with $R^2 = 0.40$ (compared to 0.32 for the NDVI and the best NDSI). It required nine bands and five pairs to produce the best linear regression model (Table 3).

Table 3. The multivariable linear regression (MLR) model performance for the highest-scoring NDSI combinations. Asterisks (*) represent significant correlations at $p < 0.001$. Highlighted in bold are the best scores (highest R^2 and lowest RMSE).

NDSI (Band1/Band2) Combinations	R2	RMSE
680/750	0.33 *	0.73
680/750, 530/623	0.35 *	0.72
680/750, 530/623, 660/940	0.35 *	0.72
680/750, 530/623, 660/940, 519/750	0.35 *	0.72
680/750, 530/623, 660/940, 519/750, 605/709	0.40 *	0.70

4.2. Machine Learning Models

The nonlinear ML models (RF, SVM, and XGB) were not better than MLR, but the performance improved when species was included as a feature (Table 4). The best model using species as a feature was SVM, with an R^2 of 0.53 and an RPD of 1.47.

Table 4. Machine learning model performance for each algorithm and the averaged model using the three algorithms without and with species as a feature in the model. NDSI combinations of 680/750, 530/623, 605/709, 519/750, and 660/940 were used as input. All models were significant at $p < 0.001$. In bold are the best statistical scores (highest R^2 and RPD, and lowest MAE, RMSE).

Model	Without Species				With Species			
	<i>R</i> ²	MAE	RMSE	RPD	<i>R</i> ²	MAE	RMSE	RPD
RF	0.40	0.59	0.74	1.30	0.42	0.59	0.74	1.32
SVM	0.35	0.58	0.77	1.25	0.53	0.50	0.67	1.47
XGB	0.40	0.58	0.74	1.30	0.47	0.57	0.71	1.38
Averaged model	0.41	0.58	0.74	1.31	0.52	0.52	0.67	1.46

All four models (MLR and the three ML algorithms) performed the best with similar NDSI band combinations. However, the importance of each NDSI combination was not the same, depending on whether the model included species (Figure 6) or did not include species as a feature (Figure 7).

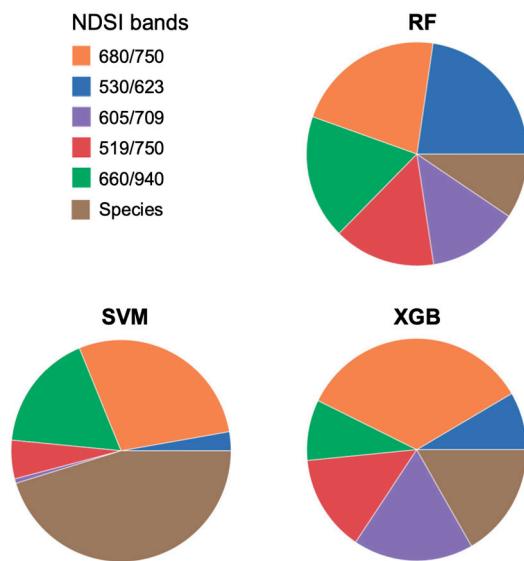


Figure 6. Feature importance of the NDSI pair bands in the different ML models when species was included as a feature.

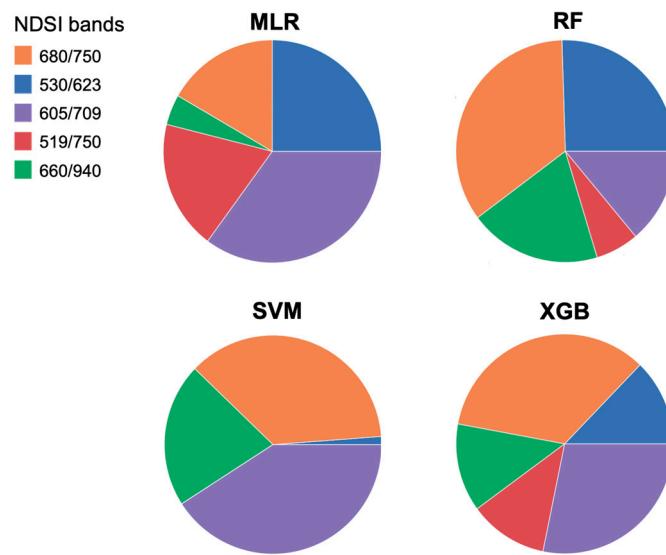


Figure 7. Feature importance of the NDSI pair bands in the MLR model and the different ML models when species were not included in the models.

Finally, the best ML model (SVM with species as a feature) predicted better ψ_{leaf} for Cypress and Pistacia than for the other species and had, specifically, low R^2 for Pine and Carob (Figure 8).

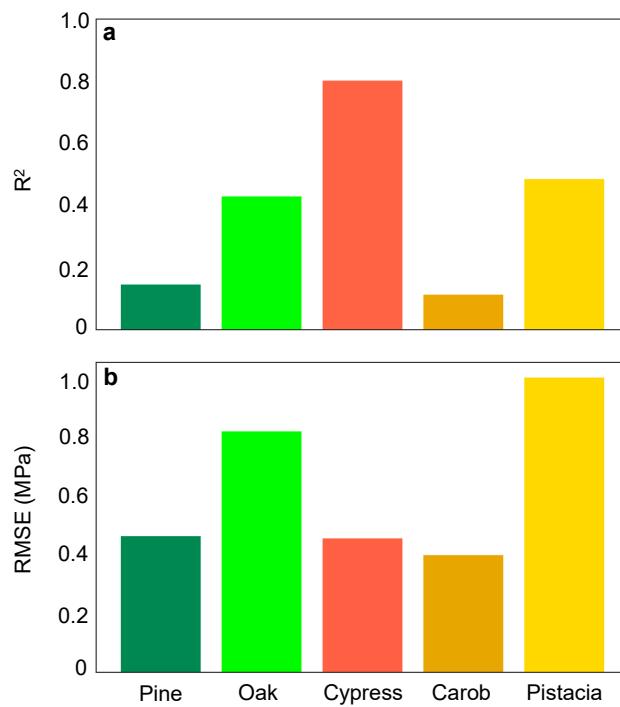


Figure 8. Performance of SVM model per species: (a) R^2 of the correlation, (b) RMSE in MPa.

Table 5 summarizes the first and second-best models for each species. It is noticeable that a simple model based on a spectral index (mostly NDVI and NDSI) could predict ψ_{leaf} and that ML models did not significantly surpass the performance of the simple single-basis linear model while requiring much more spectral information. In addition, only Cypress reached a relatively high R^2 in such models.

Table 5. A summary of the best and the second-best models for predicting ψ_{leaf} per species and for all species together (All). ML models are highlighted in gray, and spectral index models in green (light green for NDSI and dark green for vegetation index).

	Pine	Oak	Cypress	Carob	Pistacia	All
Best model	NDVI	NDVI NDSI	SVW (w/species)	NDSI	SVW (w/species)	SVW (w/species)
R^2	0.31	0.48	0.80	0.29	0.49	0.53
2nd best model	NDSI	MNDVI 940/960/NDVI	NDVI NDSI	NDVI	NDSI	Avg ML (w/o species)
R^2	0.30	0.42	0.74	0.28	0.44	0.41
Difference	0.01	0.06	0.06	0.01	0.05	0.12

■ ML model; ■ Spectral index.

4.3. Averaging Data at the Plot Scale

Due to the relatively moderate performance of the above species-specific models, we tested a ‘general’ model by averaging measurements over the entire plot (i.e., including all five species together). We found that NDSI with a band combination of 816 and 712 nm was the best predictor for ψ_{leaf} at the plot scale (Supplementary Figure S2).

Figure 9 shows observed versus predicted ψ_{leaf} using this ‘general’ model. The R^2 of the correlation was 0.79, with a relative RMSE of 13%, significant at $p < 0.001$.

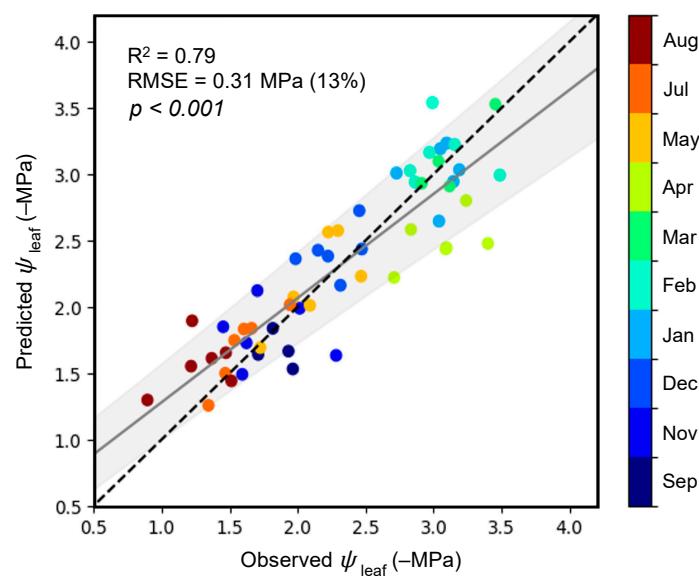


Figure 9. Predicted vs. observed ψ_{leaf} for plot-aggregated data. Each dot represents the mean value for all the species in a plot on a single date. Colors mark the months of the observed/predicted value. The broken line represents a 1:1 line.

Figure 10, showing the modeled ψ_{leaf} using the ‘general’ model, highlights the differences in ψ_{leaf} among the species (Pine with the lowest ψ_{leaf} , most negative values, Cypress with moderately negative ψ_{leaf} , and Carob with the least negative ψ_{leaf} values) and between the seasons (more negative values in September, at the end of the dry season, than in May, at the end of the rainy season).

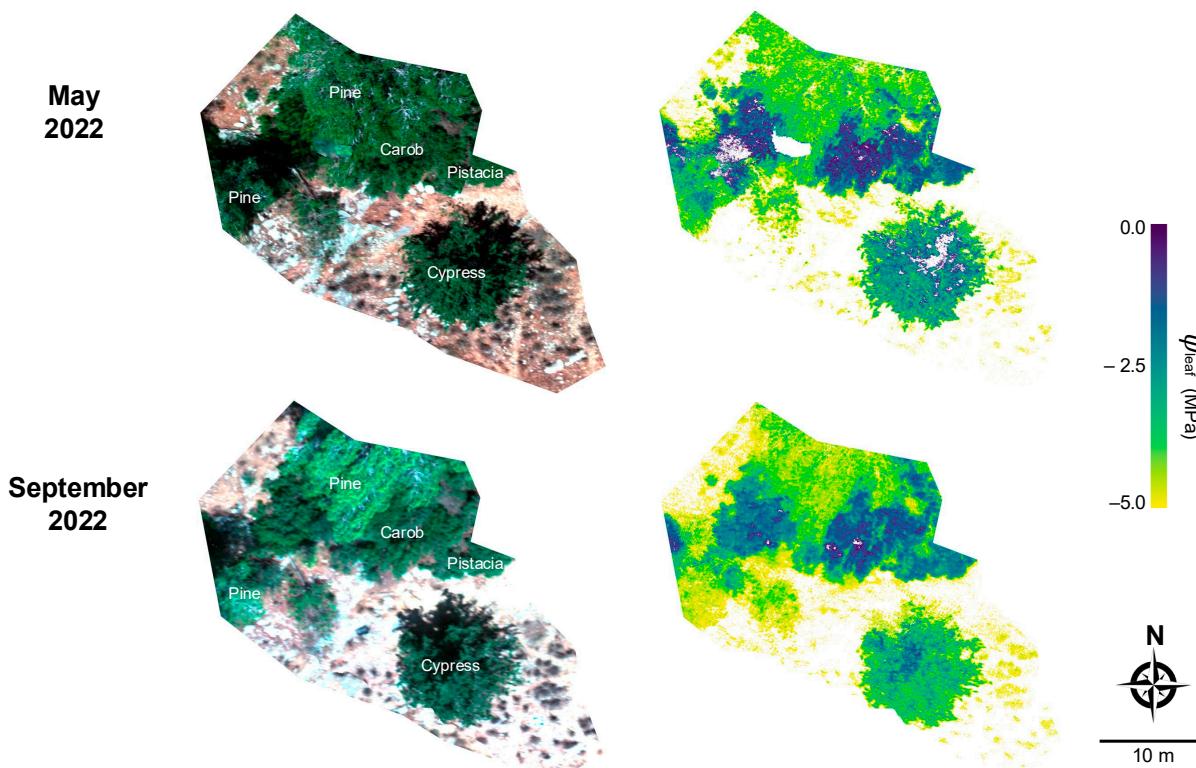


Figure 10. The ψ_{leaf} map generated using the ‘general’ model and drone images acquired on two dates, one at the end of the dry season (September) and another at the end of the wet season (May).

5. Discussion

This study explored the use of hyperspectral imaging and machine learning models to estimate leaf water potential (ψ_{leaf}) in a mixed Mediterranean forest at a high spatial resolution using drone imagery. We specifically hypothesized that ψ_{leaf} would be easier to model from the spectral data for species with a broader range of ψ_{leaf} values. This proved true since most models were much better for Cypress and Pistacia, which showed a more comprehensive range of ψ_{leaf} values along the season than the other species. We also hypothesized that machine learning algorithms should significantly enhance the ability to predict ψ_{leaf} compared to simple linear regression models using spectral indices, specifically the nonlinear models that account for more complex relationships. Indeed, our findings show that the SVM model was slightly superior to the simpler regression models using the spectral indices (Tables 2 and 4). However, except for Cypress and Pistacia, SVM was not better than simple spectral indices in predicting ψ_{leaf} (Table 5). Overall, simple linear regression with commonly used vegetation indices, like the NDVI, performed well enough to make the complex and data-demanding SVM model superfluous. Yet, the general performance of SVM above RF and XGB may be related to its ability to learn effectively with small sample datasets, thus taking advantage of the sensitivity of its support vectors to all sample data points, including outliers [61]. Also, SVM can map the dataset into high dimensional input space, where it looks for core differences between classes using the optimal hyperplane, which might have given it an added advantage when species were added as part of the feature, leading to an increased R^2 of 0.35 to 0.53 (Table 4).

The observed variations in ψ_{leaf} among the five species (Figure 5) align with previous findings by Peñuelas et al. [32], reporting distinct water management characteristics in mixed Mediterranean forest species. This variability underscores the complexity of modeling water status in heterogeneous forest ecosystems. Pine exhibited the lowest average ψ_{leaf} (-2.56 MPa) and the least inter-annual variance. In contrast, Cypress and Pistacia demonstrated the most considerable ψ_{leaf} variance, reaching as low as -4.5 MPa during dry months. It is not surprising, then, that the model performance was much better for these two species. These species-specific differences in water potential dynamics highlight the importance of considering individual species' physiological responses when developing forest water status assessment models or at least considering stand density and species as factors.

As stated, most spectral indices showed a low correlation with ψ_{leaf} (Table 2), with the NDVI being the best predictor, particularly for Cypress ($R^2 = 0.74$). The other spectral indices tested in this study exhibited suboptimal performance despite their established correlation with plant water status in previous studies. This discrepancy may be attributed to the predominant focus of these indices on leaf water content rather than ψ_{leaf} [22,44,50,56]. While the leaf water content directly influences spectral reflectance, ψ_{leaf} is a physiological parameter that reflects plant functionality and can vary in response to various factors, including water stress and changes in the ambient radiation environment. Elsayed et al. [55] demonstrated the distinction between these parameters by developing indices specifically for estimating ψ_{leaf} , independent of the leaf water content. However, the disparity between their experimental design (controlled climate chamber, handheld spectrometer, crop plants) and ours (forest canopies, hyperspectral camera at 60 m height) may explain the limited applicability of their indices in our study. Our results suggest that these indices may not translate directly to forest stands under natural conditions, where factors such as canopy structure, background soil reflectance, and atmospheric effects can influence spectral measurements.

Though using different bands, our findings align more closely with those of Stimson et al. [24], who studied two conifer species (*Pinus edulis* and *Juniperus monosperma*) in a forest

setting. They found that indices incorporating 980 nm and 1200 nm wavelength bands were significantly correlated with measured ψ_{leaf} , but the relationships differed between the two species. This species-specific variability is consistent with our observations, where Cypress and Pistacia showed stronger correlations with spectral indices than Pine and Carob. The high variability in ψ_{leaf} between species is also consistent with recent findings by Italiano et al. [25], who reported significant differences in hydraulic traits among Mediterranean forest species.

The evaluation of machine learning models, including RF, XGB, and SVM, resulted in the SVM being the best model, particularly when incorporating species as a feature (Table 4). While Sadiq et al. [30] and Virnodkar et al. [31] focused on other vegetation characteristics rather than ψ_{leaf} specifically, their work demonstrates the broader potential of machine learning in remote sensing applications for forest assessment. In our study, the performance varied across species, with Cypress and Pistacia showing better predictability compared to Pine, Oak, and Carob (Figure 8). This variation may be attributed to the higher inter-annual variance of ψ_{leaf} in Cypress and Pistacia, enabling stronger correlations between changes in leaf water potential and spectral indices. The improved performance of models incorporating species information highlights the importance of considering species-specific traits in mixed forest ecosystems. However, the practical application of such models in mixed forests remains challenging due to the difficulty in differentiating species in the image, especially when considering the most practical remote sensing use of satellites, which usually have a coarser spatial resolution. However, efforts seem successful in satellite and drone-based spectral differentiation of vegetation classes and types (e.g., [66,67]).

Our ‘general’ model, averaging all species at the plot scale, has a greater potential for satellite remote sensing use. This is because it indicates that species differentiation is not required. This ‘general’ model is based on two bands, one within the red-edge range (712 nm), which marks the inflection point from strong absorption by leaf chlorophyll in the visible region to structurally dominated reflectance in the NIR, which was the second band (816 nm) [68]. This is not surprising since band centering within the red-edge region is very sensitive to the slightest changes in chlorophyll content due to stress or LAI changes, resulting in either a blue shift—toward shorter wavelengths or a red shift—toward longer wavelengths [69–71]. The red and blue bands readily saturate at relatively low chlorophyll content or show an almost insignificant response to even small biochemical changes. Thus, the red-edge band is more likely to respond to changes in pigmentation and leaf cell or canopy structure. In contrast, bands within the visible regions are less sensitive to structure [72,73].

The plot-scale approach may offer a more robust method for assessing forest water status over larger areas, potentially bridging the gap between individual leaf measurements and landscape-scale assessments using satellite imagery. The improved performance at the plot scale has important implications for forest management and ecological monitoring. It suggests that while individual tree-level predictions may be challenging due to high variability, aggregating data at a broader scale can provide reliable estimates of forest water status. This approach could be convenient for assessing drought stress across forest stands or informing management decisions in planted forests, which comprise about half of the forested areas in Israel [35].

While promising, this ‘general’ model has two main limitations. First, it was generated at the plot scale, and the scalability to a finer scale, as shown in Figure 10, still needs to be verified. Nevertheless, the seasonal differences with more negative ψ_{leaf} observed at the end of the dry season than at the end of the wet season indicate that the model can reasonably capture temporal variations in forest water status. This temporal sensitivity is crucial for monitoring the impacts of seasonal drought and climate change on forest ecosystems, as

highlighted by Allen et al. [1] and reported by IPCC [2], even though the magnitude might be inaccurate. Moreover, the clear differentiation between species in the modeled ψ_{leaf} , in accordance with the measured values, further strengthens our confidence in the model's ability to distinguish between species-specific water management strategies. The second limitation concerns the generality of the model. In other words, how reliable can this model be when transferred to another area? Since plots were relatively homogeneous in their species distribution, the model might need to be adjusted with, at least, a factor considering the stand density (when using satellite imagery) or/and by considering different species mixing at the pixel level. However, considering that the species in our study area are common Mediterranean species and the climate is a typical Mediterranean one, our model may be applied to similar regions elsewhere. Nevertheless, further testing should be conducted in different environments to ensure the model's general applicability.

Finally, ψ_{leaf} is not a trivial parameter to detect directly and remotely. Without the invention of the pressure chamber and applying very high pressure (>50 bar in this study), it would have been very difficult to predict otherwise. Non-destructive ψ_{leaf} determination in the field by a few methods is still being tested. Thus, though challenging, our remote sensing effort to predict ψ_{leaf} proved successful and can be further used to detect ψ_{leaf} changes across mixed forest stands. Future research should include drought-stressed plots, which are the target of a future stress alert method.

6. Summary and Conclusions

Analysis applying vegetation indices and various ML models to the entire dataset without incorporating species information resulted in weak to moderate correlations. The absence of a significant advantage for more complex ML models can be explained by the intrinsic variability in leaf spectral features across different species within the forest. When species information was included as a feature, the correlation improved for SVM but was still weak to be superior to the simple spectral index model.

A 'general' plot scale model without species-specific differentiation performed much better. This finding is promising since the model can be used with satellite images to derive ψ_{leaf} at a coarser yet continuous spatial resolution. The fact that a 'general' model was better than all species-specific models highlights the inherent challenges of applying accurate remote sensing techniques in mixed-species forests, where interspecies variability is significant. These results suggest that plot-level aggregation may offer a more robust approach for assessing forest water status over larger areas. Still, such a conclusion must be further tested since the plots in our study area were relatively homogeneous regarding species distribution and density. In cases where fine-scale evaluation is required, and species classification is available, our findings indicate that the SVM model can be relatively effective in addressing these challenges.

Our approach facilitates the precise monitoring of ψ_{leaf} , which can inform targeted forest conservation efforts, adaptive management strategies, and interventions tailored to species-specific drought vulnerabilities, thereby enhancing ecosystem resilience in the face of increasing drought stress.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs17010106/s1>; Supplementary Table S1. Optimized hyperparameters for Random Forest (RF), Support Vector Ma-chine (SVM), and XGBoost (XGB); Supplementary Figure S1. Correlation maps of all possible two-band combinations with normalized difference vegetation index (NDSI) and the measured ψ_{leaf} ($N = 246$); Supplementary Figure S2. Correlation maps of all possible two-band combinations with normalized difference vegetation index (NDSI) and the measured ψ_{leaf} averaged over the entire plot ($N = 56$).

Author Contributions: Conceptualization, D.H. and T.K.; Methodology, N.F., A.Y., S.O., L.R., Y.M., T.K. and D.H.; Software, N.F., Y.Y., G.M., Y.M. and D.H.; Validation, N.F., Y.Y., T.K. and D.H.; Formal analysis, N.F. and D.H.; Investigation, N.F., S.O., T.K. and D.H.; Resources, T.K. and D.H.; Data curation, N.F., A.Y., S.O., L.R., G.M. and T.K.; Writing—original draft, N.F. and D.H.; Writing—review & editing, N.F., Y.Y., G.M., T.K. and D.H.; Supervision, D.H.; Project administration, D.H. and T.K.; Funding acquisition, D.H. and T.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Jewish National Fund (KKL) via KKL chief scientist grant #2145 (430/21). N.F., Y.Y., and G.M. were partly funded by this grant.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: We thank Gil Lerner for flying the drone and assisting with fieldwork and Itzik Ben-Shabbat from IDO for assisting with the drone and flight permits. We also thank the Jewish National Fund (KKL) for supporting this study via KKL chief scientist grant #2145 (430/21). N.F. is an M.Sc. student at the Faculty of Agriculture, Food, and Environment under the supervision of D.H.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Allen, C.D.; Macalady, A.K.; Chenchouni, H.; Bachelet, D.; McDowell, N.; Vennetier, M.; Kitzberger, T.; Rigling, A.; Breshears, D.D.; Hogg, E.H.; et al. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For. Ecol. Manag.* **2010**, *259*, 660–684. [[CrossRef](#)]
- IPCC. *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Core Writing Team, Lee, H., Romero, J., Eds.; IPCC: Geneva, Switzerland, 2023.
- Klein, T.; Cahanovitc, R.; Sprintsin, M.; Herr, N.; Schiller, G. A nation-wide analysis of tree mortality under climate change: Forest loss and its causes in Israel 1948–2017. *For. Ecol. Manag.* **2019**, *432*, 840–849. [[CrossRef](#)]
- Patsiou, T.S.; Shestakova, T.A.; Klein, T.; di Matteo, G.; Sbay, H.; Chambel, M.R.; Zas, R.; Voltas, J. Intraspecific responses to climate reveal nonintuitive warming impacts on a widespread thermophilic conifer. *New Phytol.* **2020**, *228*, 525–540. [[CrossRef](#)]
- Hartmann, H.; Moura, C.F.; Anderegg, W.R.L.; Ruehr, N.K.; Salmon, Y.; Allen, C.D.; Arndt, S.K.; Breshears, D.D.; Davi, H.; Galbraith, D.; et al. Research frontiers for improving our understanding of drought-induced tree and forest mortality. *New Phytol.* **2018**, *218*, 15–28. [[CrossRef](#)] [[PubMed](#)]
- Jarvis, P.G. The Interpretation of the Variations in Leaf Water Potential and Stomatal Conductance Found in Canopies in the Field. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **1976**, *273*, 593–610.
- Sack, L.; Ball, M.; Brodersen, C.; Davis, S.; Des Marais, D.; Donovan, L.; Givnish, T.; Hacke, U.; Huxman, T.; Jansen, S.; et al. Plant hydraulics as a central hub integrating plant and ecosystem function: Meeting report for “Emerging Frontiers in Plant Hydraulics” (Washington, DC, May 2015). *Plant. Cell Environ.* **2016**, *39*, 2085–2094. [[CrossRef](#)] [[PubMed](#)]
- Scholander, P.F.; Hammel, H.T.; Bradstreet, E.D.; Hemmingsen, E.A. Sap pressure in vascular plants. *Science* **1965**, *148*, 339–346. [[CrossRef](#)] [[PubMed](#)]
- Kerr, J.T.; Ostrovsky, M. From space to species: Ecological applications for remote sensing. *Trends Ecol. Evol.* **2003**, *18*, 299–305. [[CrossRef](#)]
- Konings, A.G.; Rao, K.; Steele-Dunne, S.C. Macro to micro: Microwave remote sensing of plant water content for physiology and ecology. *New Phytol.* **2019**, *223*, 1166–1172. [[CrossRef](#)]
- Asner, G.P.; Martin, R.E.; Anderson, C.B.; Kryston, K.; Vaughn, N.; Knapp, D.E.; Bentley, L.P.; Shenkin, A.; Salinas, N.; Sinca, F.; et al. Scale dependence of canopy trait distributions along a tropical forest elevation gradient. *New Phytol.* **2017**, *214*, 973–988. [[CrossRef](#)]
- Zarco-Tejada, P.J.; Rueda, C.A.; Ustin, S.L. Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sens. Environ.* **2003**, *85*, 109–124. [[CrossRef](#)]
- Ullah, S.; Skidmore, A.K.; Naeem, M.; Schlerf, M. An accurate retrieval of leaf water content from mid to thermal infrared spectra using continuous wavelet analysis. *Sci. Total Environ.* **2012**, *437*, 145–152. [[CrossRef](#)]
- Helman, D.; Yungstein, Y.; Mulero, G.; Michael, Y. High-Throughput Remote Sensing of Vertical Green Living Walls (VGWs) in Workplaces. *Remote Sens.* **2022**, *14*, 3485. [[CrossRef](#)]
- Konings, A.G.; Gentine, P. Global variations in ecosystem-scale isohydricity. *Glob. Change Biol.* **2017**, *23*, 891–905. [[CrossRef](#)] [[PubMed](#)]

16. Anderson, L.O.; Malhi, Y.; Aragão, L.E.O.C.; Ladle, R.; Arai, E.; Barbier, N.; Phillips, O. Remote sensing detection of droughts in Amazonian forest canopies. *New Phytol.* **2010**, *187*, 733–750. [[CrossRef](#)] [[PubMed](#)]
17. Ceccato, P.; Flasse, S.; Tarantola, S.; Jacquemoud, S.; Grégoire, J.-M. Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sens. Environ.* **2001**, *77*, 22–33. [[CrossRef](#)]
18. Coops, N.C.; Stone, C.; Culvenor, D.S.; Chisholm, L.A.; Merton, R.N. Chlorophyll content in eucalypt vegetation at the leaf and canopy scales as derived from high resolution spectral data. *Tree Physiol.* **2003**, *23*, 23–31. [[CrossRef](#)] [[PubMed](#)]
19. Helman, D.; Mussery, A. Using Landsat satellites to assess the impact of check dams built across erosive gullies on vegetation rehabilitation. *Sci. Total Environ.* **2020**, *730*, 138873. [[CrossRef](#)] [[PubMed](#)]
20. Roberts, D.A.; Barbara, S.; Roth, K.; Perroy, R.L. *Spectral and Spatial Methods of Hyperspectral Image Analysis for Estimation of Biophysical and Biochemical Properties of Agricultural Crops*; CRC Press: Boca Raton, FL, USA, 2011; ISBN 6087858334. [[CrossRef](#)]
21. Thenkabail, P.S.; Lyon, J.G.; Huete, A. (Eds.) *Hyperspectral Indices and Image Classifications for Agriculture and Vegetation*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2018. [[CrossRef](#)]
22. Raj, R.; Walker, J.P.; Vinod, V.; Pingale, R.; Naik, B.; Jagarlapudi, A. Leaf water content estimation using top-of-canopy airborne hyperspectral data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102393. [[CrossRef](#)]
23. Peñuelas, J.; Fillla, I.; Serrano, L.; Savé, R. Cell wall elasticity and Water Index (R970 nm/R900 nm) in wheat under different nitrogen availabilities. *Int. J. Remote Sens.* **1996**, *17*, 373–382. [[CrossRef](#)]
24. Stimson, H.C.; Breshears, D.D.; Ustin, S.L.; Kefauver, S.C. Spectral sensing of foliar water conditions in two co-occurring conifer species: *Pinus edulis* and *Juniperus monosperma*. *Remote Sens. Environ.* **2005**, *96*, 108–118. [[CrossRef](#)]
25. Italiano, S.S.P.; Julio Camarero, J.; Borghetti, M.; Colangelo, M.; Pizarro, M.; Ripullone, F. Radial growth, wood anatomical traits and remote sensing indexes reflect different impacts of drought on Mediterranean forests. *For. Ecol. Manag.* **2023**, *548*, 121406. [[CrossRef](#)]
26. Othman, Y.; Steele, C.; VanLeeuwen, D.; Heerema, R.; Bawazir, S.; St. Hilaire, R. Remote sensing used to detect moisture status of pecan orchards grown in a desert environment. *Int. J. Remote Sens.* **2014**, *35*, 949–966. [[CrossRef](#)]
27. Wang, Z.; Sun, Z.; Lu, S. Optimal vegetation index for assessing leaf water potential using reflectance factors from the adaxial and abaxial surfaces. *Comput. Electron. Agric.* **2020**, *172*, 105337. [[CrossRef](#)]
28. Li, J.; Wijewardane, N.K.; Ge, Y.; Shi, Y. Improved chlorophyll and water content estimations at leaf level with a hybrid radiative transfer and machine learning model. *Comput. Electron. Agric.* **2023**, *206*, 107669. [[CrossRef](#)]
29. Li, W.; Pacheco-Labrador, J.; Migliavacca, M.; Miralles, D.; Hoek van Dijke, A.; Reichstein, M.; Forkel, M.; Zhang, W.; Frankenberg, C.; Panwar, A.; et al. Widespread and complex drought effects on vegetation physiology inferred from space. *Nat. Commun.* **2023**, *14*, 4640. [[CrossRef](#)] [[PubMed](#)]
30. Sadiq, M.A.; Sarkar, S.K.; Raisa, S.S. Meteorological drought assessment in northern Bangladesh: A machine learning-based approach considering remote sensing indices. *Ecol. Indic.* **2023**, *157*, 111233. [[CrossRef](#)]
31. Virnodkar, S.S.; Pachghare, V.K.; Patil, V.C.; Jha, S.K. Remote sensing and machine learning for crop water stress determination in various crops: A critical review. *Precis. Agric.* **2020**, *21*, 1121–1155. [[CrossRef](#)]
32. Peñuelas, J.; Sardans, J.; Filella, I.; Estiarte, M.; Llusia, J.; Ogaya, R.; Carnicer, J.; Bartrons, M.; Rivas-Ubach, A.; Grau, O.; et al. Impacts of Global Change on Mediterranean Forests and Their Services. *Forests* **2017**, *8*, 463. [[CrossRef](#)]
33. Rog, I.; Tague, C.; Jakoby, G.; Megidish, S.; Yaakobi, A.; Wagner, Y.; Klein, T. Interspecific Soil Water Partitioning as a Driver of Increased Productivity in a Diverse Mixed Mediterranean Forest. *J. Geophys. Res. Biogeosci.* **2021**, *126*, e2021JG006382. [[CrossRef](#)]
34. Lapidot, O.; Ignat, T.; Rud, R.; Rog, I.; Alchanatis, V.; Klein, T. Use of thermal imaging to detect evaporative cooling in coniferous and broadleaved tree species of the Mediterranean maquis. *Agric. For. Meteorol.* **2019**, *271*, 285–294. [[CrossRef](#)]
35. Amir, S.; Rechtman, O. The development of forest policy in Israel in the 20th century: Implications for the future. *For. Policy Econ.* **2006**, *8*, 35–51. [[CrossRef](#)]
36. Fassnacht, F.E.; Latifi, H.; Stereńczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.T.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* **2016**, *186*, 64–87. [[CrossRef](#)]
37. Pozner, E.; Bar-On, P.; Livne-Luzon, S.; Moran, U.; Tsamir-Rimon, M.; Dener, E.; Schwartz, E.; Rotenberg, E.; Tatarinov, F.; Preisler, Y.; et al. A hidden mechanism of forest loss under climate change: The role of drought in eliminating forest regeneration at the edge of its distribution. *For. Ecol. Manag.* **2022**, *506*, 119966. [[CrossRef](#)]
38. Rog, I.; Hilman, B.; Fox, H.; Yalin, D.; Qubaja, R.; Klein, T. Increased belowground tree carbon allocation in a mature mixed forest in a dry versus a wet year. *Glob. Change Biol.* **2024**, *30*, e17172. [[CrossRef](#)]
39. Israel Meteorological Service. Available online: <http://www.ims.gov.il> (accessed on 29 December 2024).
40. Sternberg, M.; Golodets, C.; Gutman, M.; Perevolotsky, A.; Kigel, J.; Henkin, Z. No precipitation legacy effects on above-ground net primary production and species diversity in grazed Mediterranean grassland: A 21-year experiment. *J. Veg. Sci.* **2017**, *28*, 260–269. [[CrossRef](#)]
41. Boyer, J.S. Leaf Water Potentials Measured with a Pressure Chamber. *Plant Physiol.* **1967**, *42*, 133–137. [[CrossRef](#)] [[PubMed](#)]

42. Ritchie, G.A.; Hinckley, T.M. The Pressure Chamber as an Instrument for Ecological Research. *Adv. Ecol. Res.* **1975**, *9*, 165–254. [[CrossRef](#)]
43. Savitzky, A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
44. Azar, M.; Mulero, G.; Oppenheimer-Shaanan, Y.; Helman, D.; Klein, T. Aboveground responses to belowground root damage detected by non-destructive sensing metrics in three tree species. *Forestry* **2023**, *96*, 672–689. [[CrossRef](#)]
45. Peñuelas, J.; Filella, I.; Biel, C.; Serrano, L.; Savé, R. The reflectance at the 950–970 nm region as an indicator of plant water status. *Int. J. Remote Sens.* **1993**, *14*, 1887–1905. [[CrossRef](#)]
46. Sadeh, R.; Avneri, A.; Tubul, Y.; Lati, R.N.; Bonfil, D.J.; Peleg, Z.; Herrmann, I. Chickpea leaf water potential estimation from ground and VENµS satellite. *Precis. Agric.* **2024**, *25*, 1–26. [[CrossRef](#)]
47. Helman, D.; Osem, Y.; Yakir, D.; Lensky, I.M. Relationships between climate, topography, water use and productivity in two key Mediterranean forest types with different water-use strategies. *Agric. For. Meteorol.* **2017**, *232*, 319–330. [[CrossRef](#)]
48. Helman, D.; Bahat, I.; Netzer, Y.; Ben-Gal, A.; Alchanatis, V.; Peeters, A.; Cohen, Y. Using Time Series of High-Resolution Planet Satellite Images to Monitor Grapevine Stem Water Potential in Commercial Vineyards. *Remote Sens.* **2018**, *10*, 1615. [[CrossRef](#)]
49. Mulero, G.; Jiang, D.; Bonfil, D.J.; Helman, D. Use of thermal imaging and the photochemical reflectance index (PRI) to detect wheat response to elevated CO₂ and drought. *Plant. Cell Environ.* **2023**, *46*, 76–92. [[CrossRef](#)] [[PubMed](#)]
50. Hunt, E.R., Jr.; Daughtry, C.; Qu, J.; Wang, L.; Hao, X. Comparison of hyperspectral retrievals with vegetation water indices for leaf and canopy water content. *Proc. SPIE Int. Soc. Opt. Eng.* **2011**, *8156*, 5. [[CrossRef](#)]
51. Gao, B. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [[CrossRef](#)]
52. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
53. Kriegler, F.J.; Malila, W.A.; Nalepka, R.F.; Richardson, W. Preprocessing Transformations and Their Effects on Multispectral Recognition. In Proceedings of the Sixth International Symposium on Remote Sensing of Environment; Arbor, A., Ed.; University of Michigan: Michigan, MI, USA, 1969; pp. 97–131.
54. Gamon, J.A.; Peñuelas, J.; Field, C.B. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [[CrossRef](#)]
55. Elsayed, S.; Mistele, B.; Schmidhalter, U. Can changes in leaf water potential be assessed spectrally? *Funct. Plant Biol.* **2011**, *38*, 523–533. [[CrossRef](#)]
56. Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [[CrossRef](#)]
57. Inoue, Y.; Peñuelas, J.; Miyata, A.; Mano, M. Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO₂ flux measurements in rice. *Remote Sens. Environ.* **2008**, *112*, 156–172. [[CrossRef](#)]
58. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
59. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016; Volume 13, pp. 785–794.
60. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
61. Kok, Z.H.; Mohamed Shariff, A.R.; Alfatni, M.S.M.; Khairunniza-Bejo, S. Support Vector Machine in Precision Agriculture: A review. *Comput. Electron. Agric.* **2021**, *191*, 106546. [[CrossRef](#)]
62. Pinhanez, C.S.; Cavalin, P.R. Exploring the Advantages of Dense-Vector to One-Hot Encoding of Intent Classes in Out-of-Scope Detection Tasks. *arXiv* **2022**, arXiv:2205.09021.
63. Agussabti; Rahmaddiansyah; Satriyo, P.; Munawar, A.A. Data analysis on near infrared spectroscopy as a part of technology adoption for cocoa farmer in Aceh Province, Indonesia. *Data Br.* **2020**, *29*, 105251. [[CrossRef](#)]
64. Sahoo, M.M.; Perach, O.; Shachter, A.; Gonda, I.; Porwal, A.; Dudai, N.; Herrmann, I. Spectral estimation of carnosic acid content in vivo rosemary plants. *Ind. Crops Prod.* **2022**, *187*, 115292. [[CrossRef](#)]
65. Mouazen, A.M.; De Baerdemaeker, J.; Ramon, H. Effect of wavelength range on the measurement accuracy of some selected soil constituents using visual-near infrared spectroscopy. *J. Near Infrared Spectrosc.* **2006**, *14*, 189–199. [[CrossRef](#)]
66. Weil, G.; Lensky, I.M.; Resheff, Y.S.; Levin, N. Using Near-Surface Observations for Optimizing the Timing of Overhead Image Acquisition for Applied Mapping of Woody Vegetation Species. In Proceedings of the IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 5398–5401.
67. Weil, G.; Lensky, I.M.; Levin, N. Using ground observations of a digital camera in the VIS-NIR range for quantifying the phenology of Mediterranean woody species. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *62*, 88–101. [[CrossRef](#)]
68. Croft, H.; Chen, J.M. Leaf Pigment Content. *Compr. Remote Sens.* **2018**, *3*, 117–142. [[CrossRef](#)]

-
- 69. Thenkabail, P.S.; Lyon, J.G.; Huete, A. Hyperspectral Remote Sensing of Vegetation. CRC Press: Boca Raton, FL, USA, 2016; ISBN 978-1-4398-4538-7.
 - 70. Xie, Q.; Dash, J.; Huang, W.; Peng, D.; Qin, Q.; Mortimer, H.; Casa, R.; Pignatti, S.; Laneve, G.; Pascucci, S.; et al. Vegetation Indices Combining the Red and Red-Edge Spectral Information for Leaf Area Index Retrieval. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1482–1492. [[CrossRef](#)]
 - 71. Guyot, G.; Baret, F.; Major, D.J. High spectral resolution: Determination of spectral shifts between the red and infrared. *Int. Arch. Photogramm. Remote Sens.* **1988**, *11*, 750–760. [[CrossRef](#)]
 - 72. Kumar, L.; Schmidt, K.; Dury, S.; Skidmore, A. *Imaging Spectrometry and Vegetation Science BT—Imaging Spectrometry: Basic Principles and Prospective Applications*; Meer, F.D., van der, J., De, S.M., Eds.; Springer: Dordrecht, The Netherlands, 2001; pp. 111–155, ISBN 978-0-306-47578-8.
 - 73. Dong, T.; Liu, J.; Shang, J.; Qian, B.; Ma, B.; Kovacs, J.M.; Walters, D.; Jiao, X.; Geng, X.; Shi, Y. Assessment of red-edge vegetation indices for crop leaf area index estimation. *Remote Sens. Environ.* **2019**, *222*, 133–143. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

גופח 3 – מאמר 2 (Preprint)

Yungstein Y, Fishman N, Lerner G, Mulero G, Michael Y, Yaakobi A, Obersteiner S, Rez L, Klein T, Helman David. **Early detection of drought-stressed stands in Mediterranean forests using machine learning classification models and a rainfall exclusion experiment.** (*Under review*).

Preprint: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5292430.

1 **Early detection of drought-stressed stands in Mediterranean forests using machine
2 learning classification models and a rainfall exclusion experiment**

3 Yehuda Yungstein^{1,2}, Netanel Fishman^{1,2}, Gil Lerner¹, Gabriel Mulero^{1,2}, Yaron Michael¹,
4 Assaf Yaakobi³, Sophie Obersteiner³, Laura Rez³, Tamir Klein³, David Helman^{1,2,*}

5 1. Department of Soil and Water Sciences, Institute of Environmental Sciences, The Robert H. Smith
6 Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot
7 7610001, Israel.

8 2. The Advanced School for Environmental Studies, The Hebrew University of Jerusalem, Givat Ram,
9 Jerusalem 9190401, Israel.

10 3. Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, 7610001,
11 Israel.

12 * Corresponding author: david.helman@mail.huji.ac.il

13 **Abstract**

14 Climate change-driven droughts increasingly threaten Mediterranean forests. Early detection
15 is crucial for mitigating long-term impacts; yet, conventional methods are limited in spatial and
16 temporal coverage. Remote sensing offers a large-scale solution, but its application at the
17 individual-tree level remains limited, particularly in mixed-species forests.

18 We combined a controlled rainfall exclusion experiment with drone-based hyperspectral
19 imaging and machine learning to classify drought stress at the individual-tree level in a semi-
20 arid Mediterranean forest (Yishi Forest, Israel). Six 0.05-ha plots with five co-occurring tree
21 species were monitored over two hydrological years. Hyperspectral data (274 bands, 400–1000
22 nm) were used as is and after synthetically simulating Planet, VENμS, and Sentinel-2 bands in
23 three machine learning classification models.

24 Results show that rainfall was reduced by nearly half in treated plots. Standard physiological
25 metrics—leaf water potential, carbon assimilation, and transpiration—showed limited
26 treatment sensitivity across most species and seasons, whereas hyperspectral-driven machine
27 learning classification models accurately distinguished between drought-treated and control
28 stands. Logistic Regression (LR) outperformed Support Vector Machines (SVM) and Random
29 Forest (RF), reaching an accuracy of 0.85, a recall of 0.94, and an F1 score of 0.83 in classifying
30 treated stands on a held-out test set. High performance persisted after reducing input to 21
31 bands. Simulated satellite spectral data showed that SVM performed best using VENμS bands
32 (accuracy = 0.74, F1 = 0.73). When applied to real VENμS imagery from three independent
33 forest sites, the model identified areas of high drought risk one to two years before visible
34 canopy decline.

35 The presented approach offers a scalable and transferable tool for real-time forest drought
36 monitoring, supporting early warning systems amid growing climate pressures.

37 *Keywords:* Drought-risk; forest decline; Remote Sensing; Support Vector Machine; rain
38 exclusion; VENμS

39 **1. Introduction**

40 Climate variability profoundly impacts forest ecosystems worldwide, with increasing
41 frequency and severity of drought events posing a significant threat to forest health and
42 productivity (Allen et al., 2010). Rising temperatures, associated with increased atmospheric
43 CO₂, elevate evaporative demand and accelerate soil drying, potentially outweighing any
44 benefits of improved water-use efficiency (Sanginés de Cárcer et al., 2018; Yuan et al., 2019).
45 Moreover, high temperatures disrupt stomatal regulation, causing rapid depletion of soil water
46 reserves. This pushes trees closer to their physiological limits—a trend expected to intensify
47 with continued global warming (Brodribb et al., 2020; Drake et al., 2018; IPCC, 2023; Urban
48 et al., 2017; Zheng et al., 2022). Mediterranean forests, characterized by dry summers and rainy
49 winters, are particularly vulnerable to these changes and are expected to face more intense and
50 prolonged dry periods (Cramer et al., 2018). The increased frequency and severity of droughts
51 in these regions can decrease forest productivity and biodiversity (Essa et al., 2023; Klein et
52 al., 2019). In addition, some tree species may be at risk of hydraulic collapse in mixed
53 Mediterranean forests during severe droughts, underscoring the critical importance of early
54 detection methods to identify vulnerable stands and initiate timely interventions (Italiano et al.,
55 2024). Detecting and monitoring drought stress in forests is crucial for understanding
56 ecosystem responses to climate change and developing effective management strategies
57 (Hartmann et al., 2018).

58 Remote sensing involves measuring electromagnetic radiation reflected or emitted from objects
59 at various wavelengths. This technique has become invaluable for monitoring forest health and
60 detecting environmental stresses over extensive spatial scales with high temporal resolution
61 (Helman et al., 2019a). These non-invasive techniques enable the assessment of vegetation
62 properties that indicate plant physiological status and stress responses (Helman et al., 2022,
63 2019b; Mulero et al., 2023; Zarco-Tejada et al., 2021). Several recent studies have
64 demonstrated the application of remote sensing for drought detection in forests, leveraging
65 various spectral bands and vegetation indices.

66 Asner et al. (2016) demonstrated the utility of airborne imaging spectroscopy for mapping
67 canopy water content and detecting drought-induced tree mortality in California forests (Asner
68 et al., 2016). Zarco-Tejada et al. (2018) used narrow-band hyperspectral indices to detect early
69 signs of water stress in olive orchards, while Hernández-Clemente et al. (2019) combined
70 visible, near-infrared, and thermal data to assess drought-induced physiological changes in
71 Mediterranean oak forests (Hernández-Clemente et al., 2014; Zarco-Tejada et al., 2018).
72 Anderson et al. (2010) used MODIS data to evaluate the 2005 Amazonian drought, revealing
73 relationships between vegetation indices and tree mortality rates (Anderson et al., 2010). Zhang
74 et al. (2017) compared various remote sensing-based drought indices across the Continental
75 United States, noting the effectiveness of vegetation-based indices like VCI for short-term
76 drought conditions (Zhang et al., 2017). Przeździecki et al. (2023) addressed the challenges of
77 applying drought indices in forests by developing a novel approach to calculate the
78 Temperature Vegetation Dryness Index (TVDI) based on temporal changes in soil moisture
79 rather than spatial heterogeneity (Przeździecki et al., 2023). In a study on Mediterranean
80 hardwood forests, Italiano et al. (2023) combined remote sensing indices with tree-ring analysis

81 and wood anatomy, revealing variability in species-specific drought responses and identifying
82 links between canopy cover, hydraulic conductivity, and growth patterns in drought-affected
83 sites (Italiano et al., 2023).

84 Recent advancements in machine learning algorithms have greatly enhanced the capacity to
85 extract meaningful information from complex remote sensing datasets (Lary et al., 2016; Li et
86 al., 2023). Several studies have demonstrated the effectiveness of various algorithms in
87 identifying drought-affected areas and assessing forest vulnerability. Olano et al. (2021) used
88 Support Vector Machines (SVM) to detect frost defoliation events in European beech forests,
89 while Mohammed et al. (2022) compared multiple algorithms for drought prediction in the
90 eastern Mediterranean (Mohammed et al., 2022; Olano et al., 2021). Cui et al. (2022) employed
91 Long Short-Term Memory models to project evapotranspiration and assess water stress
92 vulnerability in Mediterranean-climate forests. Buthelezi et al. (2022) evaluated several
93 machine-learning techniques for classifying drought-damaged trees using Landsat-derived
94 vegetation indices in subtropical regions. Additionally, machine learning approaches such as
95 random forests (RF), SVM, and artificial neural networks (ANN) have been successfully used
96 to estimate various water-related plant parameters, including leaf water content, leaf water
97 potential, and equivalent water thickness, from both hyperspectral and multispectral remote
98 sensing data (Fishman et al., 2025; Li et al., 2023; Sadiq et al., 2023; Virnodkar et al., 2020).

99 Despite the increasing use of remote sensing techniques for forest monitoring, which relies
100 mainly on retrospective analyses of natural drought events, a critical gap remains in accurately
101 classifying drought-stressed stands in forests. No study has yet combined high-resolution
102 hyperspectral drone imagery with machine learning algorithms to classify experimentally
103 induced drought stress at the individual tree level in a forest ecosystem. This approach is
104 particularly important for monitoring stands in Mediterranean forests, where the complex
105 composition of multiple tree species and their varied drought sensitivity complicate monitoring
106 efforts (Cramer et al., 2018; Klein et al., 2019).

107 Here, we combine hyperspectral drone imagery with machine learning algorithms to classify
108 drought stress in a mixed Mediterranean forest under a unique natural controlled rainfall
109 reduction experiment. We compare three machine learning classification algorithms,
110 accounting for heterogeneous responses of co-occurring species, and integrate high-resolution
111 spectral data with physiological measurements across five tree species to develop drought
112 classification models at the individual tree level. We further train new models based on
113 synthetic data from drone-based hyperspectral images based on Sentinel-2, VENμS, and Planet
114 satellite bands.

115 **2. Data**

116 *2.1. Study site and experimental design*

117 The research was conducted in Yishi Forest, a semi-arid mixed Mediterranean woodland in
118 Israel's Judean foothills ($31^{\circ} 43' N$ $34^{\circ} 57' E$; Fig. 1). This forest covers approximately 650
119 hectares and is located 4 kilometers southwest of Beit Shemesh at an average elevation of 300
120 meters above sea level (Lapidot et al., 2019; Rog et al., 2024). The climate is characterized by
121 a mean annual precipitation of 460 mm, primarily occurring between November and May,

122 based on data from the past two decades. Temperature records from the Israel Meteorological
123 Service indicate a mean annual temperature of 20.4 ± 6.8 °C, with winter (January-March) and
124 summer (June-August) averages of 16.5 ± 1.9 °C and 24.1 ± 7.9 °C, respectively (Israel
125 Meteorological Service). The predominant soil type in Yishi is terra rossa, comprising A and
126 C horizons. The C horizon soil infiltrates fissures within the weathered limestone bedrock. The
127 A horizon depth averages 21 cm, ranging from 16 to 25 cm (Rog et al., 2021).

128 The vegetation in Yishi Forest includes both planted and native Mediterranean woody species.
129 The planted gymnosperms are *Pinus halepensis* (Pine) and *Cupressus sempervirens* (Cypress),
130 while the indigenous angiosperms include *Quercus calliprinos* (Oak), *Ceratonia siliqua*
131 (Carob), and *Pistacia lentiscus* (Pistacia). These species have been previously investigated in
132 situ for their water relations and carbon management strategies (Rog et al., 2024, 2021). The
133 forest understory supports a diverse community of annual plants, which flourish during the
134 winter-spring period.

135 The rainfall exclusion experiment started in November 2021 and was conducted in six 0.05-
136 hectare plots within the forest, each containing the five co-occurring woody species (Fig. 1b,c).
137 Three plots were subjected to rainfall reduction, while three served as controls. The drought
138 simulation employed an open-pipe harvesting system with gutters (Fig. 1d,e) to divert
139 approximately 50% of incident precipitation from treated plots. Soil moisture sensors (EC-5;
140 Meter; Pullman, WA, USA) monitored treatment efficacy, aiming to reduce soil moisture
141 content by 50% compared to control plots. Sensors determined volumetric water content by
142 measuring the dielectric constant of the media using capacitance/frequency domain
143 technology. Five sensors were installed in each stand, in locations under tree canopies and
144 between trees, in undisturbed soil volumes at depths of 15-20 cm below the surface. In each
145 stand, the five sensors were connected to a datalogger (ZL6; Meter; Pullman, WA, USA),
146 which recorded measurements at an hourly resolution and was downloaded during field
147 measurement days.

148 [Figure 1]

149 2.2. Field measurements

150 To monitor trees' physiological response to rainfall reduction, three key parameters were
151 assessed: leaf water potential (ψ_{leaf}), assimilation rate (A_n), and transpiration rate (T_r). These
152 measurements provided insights into the trees' water relations status and photosynthetic activity
153 under varying rainfall conditions (Blackman et al., 2009; Flexas et al., 2004).

154 ψ_{leaf} was measured using the pressure chamber method (Boyer, 1967). This technique involves
155 sealing a leaf petiole within a chamber and incrementally raising the internal pressure until sap
156 emerges from the cut end of the petiole. The pressure at this point equals the negative of the
157 ψ_{leaf} , which directly measures the leaf's hydration status (Ritchie and Hinckley, 1975).

158 ψ_{leaf} measurements were conducted monthly from October 2021 to March 2023. Sampling was
159 conducted between 11 AM and 1 PM to capture peak daily ψ_{leaf} values, with the specific time
160 adjusted seasonally. In each plot, one leaf per species was sampled to ensure consistent
161 conditions across species. To minimize measurement errors due to time lags, excised leaves

162 were immediately sealed in airtight plastic bags and kept cool. For analysis, ~30 leaves were
163 sampled per measurement date using a PMS1515 pressure chamber (PMS, Albany, OR, USA).
164 In total, 480 leaf samples were collected over the study period.

165 In addition to ψ_{leaf} , gas exchange parameters were measured to assess the trees' physiological
166 responses to the rainfall reduction treatment. A_n and T_r were measured monthly from October
167 2021 to March 2023 on clear sky days using a portable infrared gas analyzer system (IRGA;
168 GFS-3000, Walz). Measurements were conducted on mature leaves concurrently with ψ_{leaf}
169 assessments, allowing for non-destructive, in situ evaluation of gas exchange dynamics.

170 The GFS-3000 was configured with the following settings: standard leaf chamber (Walz 3010-S),
171 an ambient CO₂ concentration of 400 ppm, a flow rate of 750 $\mu\text{mol s}^{-1}$, and an impeller
172 speed of 7 steps. The temperature was set to ambient with a 1°C offset. The projected leaf area
173 relative to the chamber size was calculated and adjusted for each tree species to ensure accurate
174 measurements.

175 A_n , representing CO₂ uptake by the leaves, was measured in $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$. T_r , indicating
176 leaf transpiration, was recorded in $\text{mmol H}_2\text{O m}^{-2} \text{ s}^{-1}$. These measurements provided insights
177 into plant gas exchange dynamics under varying environmental conditions and treatments.

178 Statistical analyses were performed to examine the effects of the drought treatment on tree
179 physiology. For each measured parameter, t-tests were conducted on control and drought-
180 treated plots within each season (summer: June-August, autumn: September-November,
181 winter: December-February, spring: March-May) to quantify treatment effects on tree
182 physiological responses across seasonally varying conditions.

183 *2.3. Hyperspectral data acquisition and preprocessing*

184 Hyperspectral imagery was collected using a Nano-Hyperspec camera (Headwall Photonics)
185 mounted on a DJI Matrice 600 Pro (M600) Hexacopter. The M600, equipped with a Global
186 Navigation Satellite System (GNSS) GPS and an Inertial Measurement Unit (IMU), was
187 operated via a remote-control transmitter and a ground control station. The NanoSpec sensor,
188 a push-broom hyperspectral device, captured 274 spectral bands across 640 spatial pixels
189 within the 400-1000 nm range.

190 Image acquisition coincided with leaf measurements between 11:00 AM and 1:00 PM. The
191 drone flew 60 m above ground level, yielding a spatial resolution of 2 – 3 cm per pixel. Three
192 flights were required to cover the entire study area. Radiometric calibration employed an in-
193 situ 3x3 m grey-white reflectance panel with three distinct reflectance factors (56%, 30%, and
194 11%). Geometric corrections were based on a ground GNSS receiver (Trimble SPS585
195 precision RTK) to collect static geolocation data, enabling post-processing kinematic (PPK)
196 flight trajectory calculations.

197 Raw hyperspectral image cubes underwent radiometric calibration, geometric corrections, and
198 ortho-mosaicking using SpectralView software (version 3.1.4, Headwall Photonics). A two-
199 stage masking process removed non-representative pixels (Fig. 2). First, pixels with a
200 Normalized Difference Vegetation Index (NDVI; Rouse, 1973) value lower than 0.3 were
201 excluded to eliminate soil and understory vegetation (Fig. 1a). Second, shaded canopy portions

were masked using a near-infrared reflectance threshold below 0.07 – 0.2 (Fig. 1b). These NDVI and NIR thresholds were determined through iterative visual inspection to optimize removing extraneous elements while retaining vegetation pixels (Fig. 1c). The masking process was done using the Quantum GIS free software (version- 3.32.3).

[Figure 2]

For each tree, the reflectance values of the remaining pixels were averaged to obtain a single mean reflectance spectrum representative of the entire canopy. To mitigate potential artifacts and anomalies in the spectral signatures, a Savitzky-Golay filter (Savitzky and Golay, 1964) was applied, using a window size of 20 bands and second-order polynomials.

The sampling strategy yielded 25 –30 spectral samples per species for each treatment, with one exception. Due to its understory growth habit, *Pistacia* yielded only four spectral samples from drought-treated plots compared to 22 from control plots. In total, the study comprised 246 samples, consisting of 125 from control plots and 121 from the rainfall reduction treatments.

2.4. Synthetic and actual satellite data

To enhance the approach's applicability, we used the bands of three high-resolution satellites: Sentinel-2, VENμS, and Planet. We synthetically produced the satellite-equivalent bands from the drone's hyperspectral images, which fall within the same 400–1000 nm range as our Headwall Photonics' hyperspectral camera (Table 1).

[Table 1]

Synthetic bands were created for each satellite by averaging the hyperspectral data over the wavelengths corresponding to each satellite band to train and evaluate the models. We further used the model with actual satellite data (for the optimal satellite/model combination) at three additional Mediterranean forest sites along the rainfall gradient in Israel (see Section 3.3 below), providing insights into the model's ability to detect early drought-stressed stands using real satellite data.

Following the results of the synthetic model evaluation, we obtained actual satellite imagery of the best satellite platform for the three case study sites (Tzora Forest (Tz), Shacharia Forest (Sh), and Gilboa Forest (G) shown in Fig. 1a). Data was downloaded from the Israel VENμS data portal, maintained by Ben-Gurion University of the Negev (<https://venus.bgu.ac.il/venus/>) for: Tz site, dates 11/9/18, 19/9/19, 19/9/20, site G – 1/12/17, 1/9/2018, 4/9/20, and site Sh – 1/9/18, 2/9/19, 4/9/20. Images were obtained from identical months across multiple years to control for phenological and seasonal variability, spanning three key temporal phases: pre-drought conditions, drought year, and post-drought management intervention periods as documented by JNF (see Section 3.3 and Fig. S1). Level-2 products were downloaded, providing surface reflectance after atmospheric correction for single-day acquisitions at a spatial resolution of 5 m. The data were provided in ready-to-use GeoTIFF format, and no additional preprocessing was required prior to analysis.

3. Methods

240 *3.1. Machine learning classification models*

241 We tested three machine learning classification algorithms to build a model that distinguishes
242 drought from control stands using only hyperspectral data: Logistic Regression (LR), Support
243 Vector Machine (SVM), and Random Forest (RF). LR estimates the probability of an outcome
244 using a linear model based on input variables (Cox, 1958), SVM identifies the hyperplane that
245 best separates classes in a high-dimensional space (Cortes and Vapnik, 1995), and RF
246 constructs multiple decision trees and assigns the class that is the mode of the classes predicted
247 by these trees (Breiman, 2001).

248 Input features comprised average canopy reflectance values across 274 hyperspectral bands,
249 for the hyperspectral data models, and fewer bands, for the satellite synthetic data model (see
250 **Table 1**). The predicted variable consisted of 246 samples, comprising 125 controls and 121
251 drought samples, which were randomly partitioned into a training set (70%, N=172) and a held-
252 out test set (30%, N=74), with reproducibility ensured through a fixed random state.

253 *3.1.1. Hyperparameter selection*

254 Model hyperparameters were systematically optimized to enhance performance and mitigate
255 overfitting, using a randomized search strategy coupled with cross-validation restricted to the
256 training dataset.

257 For LR, optimization employed a 3-fold cross-validation scheme across 50 iterations, with
258 accuracy as the optimization metric. The hyperparameter space included penalty type, inverse
259 regularization strength, solver algorithm, convergence tolerance, intercept inclusion, and
260 intercept scaling factor (applicable only with specific solver and intercept configurations), as
261 well as maximum iterations ranging from 100 to 500.

262 For SVM, accuracy optimization used a 3-fold cross-validation across 50 iterations. The
263 hyperparameter space encompassed the following options: kernel type (linear, polynomial,
264 radial basis function, and sigmoid), shrinking heuristic utilization, regularization parameter,
265 kernel coefficients relevant for polynomial and sigmoid kernels, polynomial degree for
266 polynomial kernels, and convergence tolerance.

267 For RF, hyperparameter optimization employed a 3-fold cross-validation scheme across 50
268 iterations, with accuracy as the optimization metric. Optimized parameters included the number
269 of trees ranging from 200 to 2000, a split quality criterion based on either Gini impurity or
270 entropy, maximum tree depth, minimum samples required for node splitting, minimum samples
271 required per leaf node, and the number of features considered for optimal splitting. The search
272 used all available processor cores.

273 *3.1.2. Feature selection*

274 We implemented a multi-step approach to address potential overfitting due to the high
275 dimensionality of the data for models using all 274 spectral bands (for only 246 samples), as
276 described below. Models using the satellite synthetic data did not require dimensionality
277 reduction.

278 To facilitate this dimensionality reduction, feature importance was quantified for each model
279 trained on the training set ($N = 172$). The method for calculating importance varied by
280 algorithm:

281 For the LR model, importance was calculated as the absolute value of the fitted coefficients,
282 normalized to represent the percentage contribution of each spectral band. For the RF model,
283 the intrinsic mean decrease in impurity, also known as Gini importance, was calculated during
284 training and normalized to percentages. For the SVM, which used the optimized kernel and
285 required probability estimates, feature importance was estimated using Permutation
286 Importance. This involved measuring the mean decrease in model accuracy on the training data
287 when the values of individual features were randomly permuted across 10 repeats. The
288 resulting mean importance scores were normalized to percentages.

289 Feature selection was then performed for each model. Features accounting for 80% of the
290 cumulative importance were retained. If the number of selected features exceeded 25 bands
291 (10% of the sample size), an additional filter was applied to retain only the most significant
292 band within each 10 nm range. Following feature selection, models were retrained using this
293 reduced feature set, with random hyperparameter search and k-fold cross-validation applied
294 again.

295 Final evaluations of these refined models were conducted on the test set using the metrics above
296 to assess their generalizability to unseen data. The final evaluation of these refined models was
297 conducted on the held-out test set using standard classification metrics: Accuracy, Recall,
298 Precision, and F1 Score, defined in Section 3.2. In addition to these metrics, the distribution of
299 predicted probabilities for the 'drought' class, generated using probability estimation on the test
300 set, was examined for each model. Histograms comparing the distributions for control versus
301 drought actual samples were plotted to assess class separability and analyze misclassification
302 patterns.

303 Figure 3 summarizes the entire modeling scheme.

304 [Fig. 3 – Modeling schemes]

305 *3.2. Statistical analyses*

306 ψ_{leaf} , An , and T_r measurements were tested for normality at $p > 0.05$ sample with the Shapiro–
307 Wilk test using the JMP 17 Pro statistical software (SAS Institute) before applying the t-test.
308 All tests for statistical significance of model performance were performed within the Python
309 pipeline at $p < 0.05$. All ML models were implemented using the scikit-learn library (Version
310 1.6.1; Pedregosa et al., 2011) within a Python (Version 3.11.12) environment.

311 Model performance was assessed using Accuracy, Recall, Precision, and F1 Score metrics.
312 These metrics are calculated as follows:

313
$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

314
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

315
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

316
$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

317 where TP denotes True Positives, TN denotes True Negatives, FP denotes False Positives, and
 318 FN denotes False Negatives. Accuracy measures the proportion of correct predictions (both
 319 true positives and true negatives) among the total number of cases examined. Recall, also
 320 known as sensitivity, quantifies the proportion of actual positive cases that were correctly
 321 identified. In our context, it represents the model's ability to identify trees under drought stress
 322 correctly. Precision measures the proportion of positive predicted positive cases, highlighting
 323 the model's ability to accurately identify drought-stressed trees. The F1 Score is the harmonic
 324 mean of precision and recall, providing a single score that balances both metrics.

325 *3.3. Case studies for model evaluation with actual satellite data*

326 To validate the efficacy of our best-performing model beyond experimental conditions, we
 327 conducted an independent assessment across three geographically distinct Mediterranean
 328 forests in Israel: Tzora Forest (Tz), Shacharia Forest (Sh), and Gilboa Forest (G) (Fig. 1a). This
 329 approach enabled evaluation of the model's transferability to actual satellite imagery for
 330 detecting naturally occurring drought stress across diverse environmental gradients. The
 331 evaluation sites represent a range of varied microclimatic conditions within Israel's
 332 Mediterranean forest ecosystem. The Tz site has a mean annual temperature of $20.5 \pm 6.8^\circ\text{C}$
 333 and a relative humidity of $60.6 \pm 21.9\%$. The Sh site has a mean annual temperature of $20.7 \pm$
 334 7.1°C and a relative humidity of $67.0 \pm 21.0\%$. G site has a mean annual temperature of $19.1 \pm$
 335 7.0°C and a relative humidity of $66.5 \pm 23.9\%$.

336 To quantify drought occurrence at each site, we analyzed annual precipitation data from the
 337 nearest Israel Meteorological Service (IMS) stations, at distances of 5.8 km, 4.8 km, and 7.5
 338 km from Tzora, Shacharia, and Gilboa forests, respectively. 2017 saw a severe drought in all
 339 three sites (Fig. S1). Thus, we selected this year for our model evaluation and searched for
 340 signs of forest decline in subsequent years, looking for overlapping areas marked as stressed
 341 according to the model.

342 Following the comparative analysis of hyperspectral-derived synthetic satellite bands, we used
 343 the satellite platform that yielded the highest classification accuracy in our experimental setup
 344 to produce drought risk maps for the three sites (Tz, Sh, and G). The optimal classification
 345 model, as determined from our experimental site analysis, was applied directly to the raw
 346 satellite imagery. The model generated pixel-level drought probability scores ranging from 0
 347 (low drought probability) to 1 (high drought probability) across the site.

348 **4. Results**

349 *4.1. Field observations*

350 The soil water content was consistently lower at the rain-exclusion plots during the entire
 351 period (Fig. 4). The average reduction was 47.2%, implying that nearly half of the rain did not
 352 reach the ground in these plots.

353

[Figure 4]

354 However, it seems that such a reduction did not affect the ψ_{leaf} in most stands (Fig. 5), except
 355 the oak that exhibited more negative values during summer and autumn (Fig. 5c) and the pine
 356 that showed lower values during the spring (Fig. 5d). The t-tests verified that the reduction of
 357 almost half the incoming water did not affect the apparent tree physiology. At least not via
 358 ψ_{leaf} , A_n and T_r for most of the season and most species (Table 2).

359

[Figure 5]

360

[Table 2]361 *4.2. Machine learning classification models*

362 The classification models, however, showed a significant difference among the spectroscopy
 363 of drought and control stands (Table 3). The best model was the linear model (LR), with an
 364 accuracy of 0.85, a recall of 0.94, and an F1 score of 0.83. LR was the best-performing
 365 algorithm even when the number of bands was reduced to 80% of the most important bands,
 366 from 274 bands to only 21 bands.

367

[Table 3]

368 The LR was not only the best model, but it also showed to perform extremely accurate in
 369 separating drought from control stands, with only few overlapping cases (Fig. 6). Most
 370 misclassified cases were within the $\pm 20\%$ overlapping zone (i.e., between 0.3 and 0.7; Fig.
 371 S2a), with only three false positive (i.e., undetected drought stands) cases (Fig. S2b).

372

[Figure 6]373 *4.3. Satellite synthetic and actual data*

374 The best model for the satellite synthetic data was, in this case, the SVM (Table 4). Statistics
 375 show that SVM had a better performance with the bands of almost all the three satellites.
 376 However, Venus outperformed other satellites with an accuracy of 0.74 and F1 of 0.73. There
 377 was no significant difference in the importance of the bands, with B12 (900 – 920 nm) being
 378 the most important (10%), and B2 (400 – 440 nm), the least (7%; Fig. S3).

379

[Table 4]

380 Using the SVM model with the actual Venus satellite imagery data for December 2017, showed
 381 specific areas in all three sites with a high degree of drought-stress risk (red areas in Fig. 7).
 382 The RGB images of the same date, however, did not indicate forest decline or particular stands
 383 at risk. Nevertheless, the Tzora site exhibited severe tree mortality the following year (as
 384 observed in the aerial image and reported by the JNF), almost precisely in the same area
 385 predicted to be at high risk by the SVM model. The same observation was made two years later

386 at the Shacharia and Gilboa sites. Once the damage was visible (in the RGB images), the model
387 no longer indicated drought-stress risk zones (Fig. 7).

388 [Figure 7]

389 **5. Discussion**

390 Accurate canopy-level detection of incipient drought stress in Mediterranean forests stands is
391 critical for implementing timely intervention strategies in increasingly water-limited
392 environments. This study demonstrates that hyperspectral imaging coupled with advanced
393 machine learning algorithms enables the identification of water stress signatures before
394 conventional physiological metrics indicate drought conditions, representing a promising
395 advancement in precision forest monitoring methodologies.

396 The rainfall exclusion successfully reduced soil water content by approximately ~47% in the
397 treatment plots, clearly establishing differential water availability conditions. However, this
398 substantial reduction in water availability did not translate into consistent, detectable
399 differences in physiological parameters for most species and seasons (Table 2). These findings
400 align with previous studies, indicating Mediterranean woody species possess diverse adaptive
401 mechanisms to cope with seasonal water shortages (Klein et al., 2019; Lloret et al., 2004). The
402 observed physiological resilience may be attributed to: (i) leaf-level measurements
403 inadequately capturing integrated canopy responses (Zarco-Tejada et al., 2018), (ii) deep root
404 systems accessing water beyond monitored soil layers (Rog et al., 2021), and (iii) efficient
405 water-use strategies maintaining physiological function despite reduced water availability
406 (Brodribb et al., 2020; Liu et al., 2015).

407 Species-specific sensitivity was evident despite overall resilience. Oak species exhibited lower
408 ψ_{leaf} during the summer and autumn periods (Fig. 5c), while Pine showed reduced ψ_{leaf} in
409 spring (Fig. 5d; Table 2). This differential sensitivity is consistent with studies showing that
410 Mediterranean species vary in their hydraulic thresholds and drought response strategies
411 (Italiano et al., 2024; Liu et al., 2015). For instance, while the oak's response here was limited
412 to ψ_{leaf} changes, long-term experimental drought has been shown to significantly reduce
413 aboveground biomass increment in *Quercus ilex* (Liu et al., 2015), indicating that sustained
414 water deficit, even if not immediately reflected in all physiological metrics, can have
415 cumulative impacts on growth. Similarly, the observed spring sensitivity in Pine could be an
416 early indicator of this genus's known vulnerability to drought-induced mortality in the region,
417 which often becomes apparent under more severe or prolonged water stress (Klein et al., 2019).

418 Notably, despite limited detection through conventional physiological measurements, our
419 machine learning models successfully identified distinct spectral signatures associated with
420 rainfall reduction treatment using canopy-level hyperspectral data. This indicates that
421 hyperspectral imaging captures subtle, integrated changes in vegetation optical properties
422 induced by water stress, potentially reflecting biochemical changes, alterations in canopy water
423 content (Asner et al., 2016), or structural adjustments such as changes in LAI (Hernández-
424 Clemente et al., 2014) before pronounced physiological symptoms become apparent at the leaf
425 level (Zarco-Tejada et al., 2018).

426 The superior performance of the LR model (accuracy = 0.82, recall = 0.91; Table 3) indicates
427 that hyperspectral imagery effectively captures drought-induced vegetation changes, even
428 when conventional physiological measurements detect minimal differences. This linear model
429 outperformed more complex algorithms when applied to high-resolution spectral data. The
430 maintenance of high classification accuracy with only 21 bands reduced from 274 demonstrates
431 the efficiency of targeted spectral monitoring, with important practical implications for
432 operational drought detection systems. The high recall value (0.91) is particularly valuable for
433 early warning systems where missed detections could preclude timely management
434 interventions.

435 The efficacy of LR with hyperspectral data likely stems from the nature of the classification
436 problem at high spectral resolution. Hyperspectral data retain narrow absorption features that
437 create a nearly linear separation boundary between drought and control classes, allowing a
438 weighted sum of key bands to effectively distinguish between treatments (Peñuelas et al.,
439 1993). Additionally, the L1-regularized LR performs embedded feature selection, making it
440 particularly well-suited for high-dimensional data with hundreds of potentially correlated
441 bands, as evidenced by its robust performance even after dimensionality reduction.

442 The consistently suboptimal RF performance (accuracy=0.62; Table 3) can be attributed to
443 several factors: tree-split algorithms lack the global regularization necessary for high-
444 dimensional correlated data; with 246 samples distributed across 274 bands, RF requires
445 stronger signals to rise above random splits; and the limited data per terminal node leads to
446 either high-variance predictions or overly pruned trees. These findings contrast with previous
447 studies by Buthelezi et al. (2022) and Mohammed et al. (2022), which reported favorable
448 results using RF for drought classification, though in different forest ecosystems with larger
449 datasets (Buthelezi et al., 2022; Mohammed et al., 2022).

450 When extending our approach to satellite-compatible spectral resolutions, the Support Vector
451 Machine algorithm demonstrated superior performance across all three simulated satellite
452 platforms (Table 4), with VENμS satellite bands yielding the best results (accuracy = 0.74, F1
453 = 0.73). This represents a moderate reduction in performance compared to the full
454 hyperspectral dataset, but remains promising for operational applications given the wider
455 coverage and regular revisit times of satellite platforms. While broader vegetation indices
456 derived from satellites, such as the Vegetation Condition Index (VCI), are used for monitoring
457 general drought conditions across large areas, their performance can vary significantly
458 depending on the region and land cover (Zhang et al., 2017). Our approach focuses on
459 identifying physiological stress spectrally at a finer scale before such indices might show
460 significant changes.

461 The shift in optimal algorithm from LR with hyperspectral to SVM with satellite-simulated
462 data reflects a fundamental transformation in the classification problem. Aggregated satellite
463 bands blur the narrow absorption features that enable linear separation in hyperspectral data,
464 resulting in classification that now relies on non-linear interactions between bands, which
465 SVM, with its polynomial kernel, can effectively capture. The versatility of SVM for analyzing
466 forest stress using satellite data has also been demonstrated in other contexts, such as detecting
467 frost defoliation (Olano et al., 2021). RF models continued to underperform across all satellite

468 platforms (Table 4), with accuracies ranging from 0.51-0.61, substantially below both SVM
469 and LR models. Even with reduced dimensionality, RF struggled with the same fundamental
470 limitations: our relatively modest sample size ($n = 246$) provided insufficient data for stable
471 tree construction across even the reduced feature space (Barreñada et al., 2024; Han et al.,
472 2021).

473 The relative uniformity in band importance observed for the VENμS satellite, with B12 (900-
474 920 nm) being only marginally more important at 10% (Figure S3), suggests that the drought
475 signal is distributed across multiple spectral regions rather than concentrated in specific bands.

476 The model's ability to identify drought-stressed zones 1-2 years before visible forest decline
477 represents the study's most significant contribution. The SVM model applied to VENμS
478 satellite imagery successfully predicted areas of high drought-stress risk in all three test sites,
479 which subsequently experienced severe tree mortality (Fig. 7), despite showing no visual
480 symptoms in RGB imagery at the time of prediction.

481 This early detection capability advances traditional monitoring, extending pre-visual spectral
482 stress detection (Zarco-Tejada et al., 2018) to operational satellite platforms in Mediterranean
483 forests. Notably, the model identified high-risk zones before visible damage appeared but
484 ceased predictions post-decline, indicating sensitivity to active physiological stress rather than
485 advanced symptoms. This temporal specificity aligns with studies linking spectral data and
486 machine learning to physiological indicators, like earlywood hydraulics (Italiano et al., 2023)
487 or ψ_{leaf} (Fishman et al., 2025) in Mediterranean forests, and more broadly to water stress
488 assessments (Sadiq et al., 2023; Virnodkar et al., 2020).

489 The model's ability to predict drought stress in unseen areas, areas excluded from the training
490 set, indicates its generalization and robustness.

491 *5.1. Limitations and future directions*

492 Several limitations warrant consideration. First, our experimental drought simulation, although
493 substantial, may not fully replicate the complex dynamics of natural drought events, which
494 involve interactions between water limitation, heat stress, and extended duration(Allen et al.,
495 2010; IPCC, 2023). Large-scale natural droughts, such as the 2005 Amazonian event analyzed
496 using MODIS data by Anderson et al. (2010), often involve widespread, heterogeneous impacts
497 and mortality patterns that are challenging to fully replicate experimentally (Anderson et al.,
498 2010). Second, our model identified drought stress within a single growing season. Yet, longer-
499 term monitoring would enhance understanding of how spectral signatures evolve over extended
500 drought periods and seasonal cycles.

501 The performance reduction when moving from hyperspectral to multispectral satellite data
502 indicates that some drought-related spectral information is lost at coarser resolutions. Future
503 research should assess model performance over multiple years, explore its applicability in
504 diverse biomes, leverage higher spectral resolution satellite data as they become available, and
505 integrate complementary sensors, such as thermal imaging or LiDAR, to potentially enhance
506 accuracy (Jimenez-Berni et al., 2018; Przeździecki et al., 2023).

507 **6. Conclusions and implications**

508 This study demonstrates that machine learning classification models applied to hyperspectral
509 and multispectral satellite data can effectively detect early signs of drought stress in
510 Mediterranean forests, even when traditional physiological measurements fail to indicate
511 water-related stress. Our findings have several important implications for forest management
512 and conservation strategies.

513 The early detection capability we demonstrated could significantly improve the timing and
514 targeting of intervention measures, such as selective thinning or emergency irrigation,
515 potentially preventing large-scale forest dieback events. As climate change intensifies drought
516 frequency and severity in Mediterranean regions, such early warning systems become
517 increasingly valuable for preserving forest ecosystem services and biodiversity.

518 The operational implementation of our approach is facilitated by the developed Streamlit
519 application, available at <https://drought-risk-ml-analyzer.streamlit.app/>, which allows users to
520 upload VEN μ S or Sentinel-2 satellite data and receive drought risk assessments without
521 specialized remote sensing expertise. This technology transfer addresses a significant gap
522 between research advancements and practical applications.

523 **Acknowledgments**

524 We thank Gil Lerner for flying the drone and assisting with fieldwork, and Itzik Ben-Shabbat
525 from IDO for helping with the drone and flight permits. We also thank the Jewish National
526 Fund (KKL) for supporting this study via KKL Chief Scientist Grant #2145 (430/21). N.F.
527 thanks KKL for the Appleby Award. Y.Y. is a Ph.D. student, N.F. is an M.Sc. student, both
528 from the Faculty of Agriculture, Food, and Environment, Hebrew University, under the
529 supervision of D.H.

530 **References**

- 531 Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M.,
532 Kitzberger, T., Rigling, A., Breshears, D.D., Hogg, E.H. (Ted), Gonzalez, P., Fensham,
533 R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S.W., Semerci,
534 A., Cobb, N., 2010. A global overview of drought and heat-induced tree mortality reveals
535 emerging climate change risks for forests. *For. Ecol. Manage.* 259, 660–684.
536 <https://doi.org/https://doi.org/10.1016/j.foreco.2009.09.001>
- 537 Anderson, L.O., Malhi, Y., Aragão, L.E.O.C., Ladle, R., Arai, E., Barbier, N., Phillips, O.,
538 2010. Remote sensing detection of droughts in Amazonian forest canopies. *New Phytol.*
539 187, 733–750. <https://doi.org/10.1111/j.1469-8137.2010.03355.x>
- 540 Asner, G.P., Brodrick, P.G., Anderson, C.B., Vaughn, N., Knapp, D.E., Martin, R.E., 2016.
541 Progressive forest canopy water loss during the 2012–2015 California drought. *Proc. Natl.*
542 *Acad. Sci. U. S. A.* 113, E249–55. <https://doi.org/10.1073/pnas.1523397113>
- 543 Barreñada, L., Dhiman, P., Timmerman, D., Boulesteix, A.-L., Van Calster, B., 2024.
544 Understanding overfitting in random forest for probability estimation: a visualization and
545 simulation study. *Diagnostic Progn. Res.* 8, 14. <https://doi.org/10.1186/s41512-024-00177-1>
- 547 Blackman, C.J., Brodribb, T.J., Jordan, G.J., 2009. Leaf hydraulics and drought stress:
548 response, recovery and survivorship in four woody temperate plant species. *Plant, cell &*
549 *Environ.* 32 11, 1584–1595.
- 550 Boyer, J.S., 1967. Leaf Water Potentials Measured with a Pressure Chamber. *Plant Physiol.* 42,
551 133–137. <https://doi.org/10.1104/pp.42.1.133>
- 552 Breiman, L., 2001. Random Forests. *Mach. Learn.* 45, 5–32.
553 <https://doi.org/10.1023/A:1010933404324>
- 554 Brodribb, T.J., Powers, J., Cochard, H., Choat, B., 2020. Hanging by a thread? Forests and
555 drought. *Science* (80-.). 368, 261–266. <https://doi.org/10.1126/science.aat7631>
- 556 Buthelezi, M.N.M., Lottering, R.T., Hlatshwayo, S.T., Peerbhay, K.Y., 2022. Localizing the
557 analysis of drought impacts on KwaZulu-Natal's commercial forests. *Remote Sens. Appl.*
558 *Soc. Environ.* 28, 100849. <https://doi.org/https://doi.org/10.1016/j.rsase.2022.100849>
- 559 Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20, 273–297.
560 <https://doi.org/10.1007/BF00994018>
- 561 Cox, D.R., 1958. The Regression Analysis of Binary Sequences. *J. R. Stat. Soc. Ser. B* 20,
562 215–232. <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x>
- 563 Cramer, W., Guiot, J., Fader, M., Garrabou, J., Gattuso, J.-P., Iglesias, A., Lange, M.A.,
564 Lionello, P., Llasat, M.C., Paz, S., Peñuelas, J., Snoussi, M., Toreti, A., Tsimplis, M.N.,
565 Xoplaki, E., 2018. Climate change and interconnected risks to sustainable development in
566 the Mediterranean. *Nat. Clim. Chang.* 8, 972–980. <https://doi.org/10.1038/s41558-018-0299-2>
- 568 Cui, G., Ma, Q., Bales, R., 2022. Assessing multi-year-drought vulnerability in dense
569 Mediterranean-climate forests using water-balance-based indicators. *J. Hydrol.* 606,
570 127431. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2022.127431>
- 571 Drake, J.E., Tjoelker, M.G., Vårhammar, A., Medlyn, B.E., Reich, P.B., Leigh, A., Pfautsch,
572 S., Blackman, C.J., López, R., Aspinwall, M.J., Crous, K.Y., Duursma, R.A.,
573 Kumarathunge, D., De Kauwe, M.G., Jiang, M., Nicotra, A.B., Tissue, D.T., Choat, B.,

- 574 Atkin, O.K., Barton, C.V.M., 2018. Trees tolerate an extreme heatwave via sustained
575 transpirational cooling and increased leaf thermal tolerance. *Glob. Chang. Biol.* 24, 2390–
576 2402. <https://doi.org/10.1111/gcb.14037>
- 577 Essa, Y.H., Hirschi, M., Thiery, W., El-Kenawy, A.M., Yang, C., 2023. Drought characteristics
578 in Mediterranean under future climate change. *npj Clim. Atmos. Sci.* 6, 133.
579 <https://doi.org/10.1038/s41612-023-00458-4>
- 580 Fishman, N., Yungstein, Y., Yaakobi, A., Obersteiner, S., Rez, L., Mulero, G., Michael, Y.,
581 Klein, T., Helman, D., 2025. Leaf Water Potential in a Mixed Mediterranean Forest from
582 Machine Learning and Unmanned Aerial Vehicle (UAV)-Based Hyperspectral Imaging.
583 *Remote Sens.* <https://doi.org/10.3390/rs17010106>
- 584 Flexas, J., Bota, J., Loreto, F., Cornic, G., Sharkey, T.D., 2004. Diffusive and metabolic
585 limitations to photosynthesis under drought and salinity in C(3) plants. *Plant Biol.*
586 (*Stuttg.*) 6, 269–279. <https://doi.org/10.1055/s-2004-820867>
- 587 Han, S., Williamson, B.D., Fong, Y., 2021. Improving random forest predictions in small
588 datasets from two-phase sampling designs. *BMC Med. Inform. Decis. Mak.* 21, 1–9.
- 589 Hartmann, H., Moura, C.F., Anderegg, W.R.L., Ruehr, N.K., Salmon, Y., Allen, C.D., Arndt,
590 S.K., Breshears, D.D., Davi, H., Galbraith, D., Ruthrof, K.X., Wunder, J., Adams, H.D.,
591 Bloemen, J., Cailleret, M., Cobb, R., Gessler, A., Grams, T.E.E., Jansen, S., Kautz, M.,
592 Lloret, F., O'Brien, M., 2018. Research frontiers for improving our understanding of
593 drought-induced tree and forest mortality. *New Phytol.* 218, 15–28.
594 <https://doi.org/10.1111/nph.15048>
- 595 Helman, D., Bonfil, D.J., Lensky, I.M., 2019a. Crop RS-Met: A biophysical evapotranspiration
596 and root-zone soil water content model for crops based on proximal sensing and
597 meteorological data. *Agric. Water Manag.* 211, 210–219.
598 <https://doi.org/10.1016/j.agwat.2018.09.043>
- 599 Helman, D., Lensky, I.M., Bonfil, D.J., 2019b. Early prediction of wheat grain yield production
600 from root-zone soil water content at heading using Crop RS-Met. *F. Crop. Res.* 232, 11–
601 23. <https://doi.org/10.1016/j.fcr.2018.12.003>
- 602 Helman, D., Yungstein, Y., Mulero, G., Yaron, M., 2022. High-Throughput Remote Sensing
603 of Vertical Green Living Walls (VGWs) in Workplaces. *Remote Sens.* 14.
604 <https://doi.org/https://doi.org/10.3390/rs14143485>
- 605 Hernández-Clemente, R., Navarro-Cerrillo, R.M., Ramírez, F.J.R., Hornero, A., Zarco-Tejada,
606 P.J., 2014. A Novel Methodology to Estimate Single-Tree Biophysical Parameters from
607 3D Digital Imagery Compared to Aerial Laser Scanner Data. *Remote Sens.*
608 <https://doi.org/10.3390/rs61111627>
- 609 IPCC, 2023. Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II
610 and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change
611 [Core Writing Team, H. Lee and J. Romero (eds.)]. Geneva, Switzerland.
612 <https://doi.org/10.59327/IPCC/AR6-9789291691647.001>
- 613 Israel Meteorological Service, n.d.
- 614 Italiano, S.S.P., Camarero, J.J., Borghetti, M., Colangelo, M., Rita, A., Ripullone, F., 2024.
615 Drought legacies in mixed Mediterranean forests: Analysing the effects of structural
616 overshoot, functional traits and site factors. *Sci. Total Environ.* 927, 172166.
617 <https://doi.org/https://doi.org/10.1016/j.scitotenv.2024.172166>

- 618 Italiano, S.S.P., Julio Camarero, J., Borghetti, M., Colangelo, M., Pizarro, M., Ripullone, F.,
619 2023. Radial growth, wood anatomical traits and remote sensing indexes reflect different
620 impacts of drought on Mediterranean forests. *For. Ecol. Manage.* 548, 121406.
621 <https://doi.org/https://doi.org/10.1016/j.foreco.2023.121406>
- 622 Jimenez-Berni, J.A., Deery, D.M., Rozas-Larraondo, P., Condon, A. (Tony) G., Rebetzke, G.J.,
623 James, R.A., Bovill, W.D., Furbank, R.T., Sirault, X.R.R., 2018. High Throughput
624 Determination of Plant Height, Ground Cover, and Above-Ground Biomass in Wheat with
625 LiDAR. *Front. Plant Sci.* 9. <https://doi.org/10.3389/fpls.2018.00237>
- 626 Klein, T., Cahanovitc, R., Sprintsin, M., Herr, N., Schiller, G., 2019. A nation-wide analysis
627 of tree mortality under climate change: Forest loss and its causes in Israel 1948–2017. *For.*
628 *Ecol. Manage.* 432, 840–849. <https://doi.org/https://doi.org/10.1016/j.foreco.2018.10.020>
- 629 Lapidot, O., Ignat, T., Rud, R., Rog, I., Alchanatis, V., Klein, T., 2019. Use of thermal imaging
630 to detect evaporative cooling in coniferous and broadleaved tree species of the
631 Mediterranean maquis. *Agric. For. Meteorol.* 271, 285–294.
632 <https://doi.org/10.1016/j.agrformet.2019.02.014>
- 633 Lary, D.J., Alavi, A.H., Gandomi, A.H., Walker, A.L., 2016. Machine learning in geosciences
634 and remote sensing. *Geosci. Front.* 7, 3–10.
635 <https://doi.org/https://doi.org/10.1016/j.gsf.2015.07.003>
- 636 Li, J., Wijewardane, N.K., Ge, Y., Shi, Y., 2023. Improved chlorophyll and water content
637 estimations at leaf level with a hybrid radiative transfer and machine learning model.
638 *Comput. Electron. Agric.* 206, 107669.
639 <https://doi.org/https://doi.org/10.1016/j.compag.2023.107669>
- 640 Liu, D., Ogaya, R., Barbata, A., Yang, X., Peñuelas, J., 2015. Contrasting impacts of
641 continuous moderate drought and episodic severe droughts on the aboveground-biomass
642 increment and litterfall of three coexisting Mediterranean woody species. *Glob. Chang.*
643 *Biol.* 21, 4196–4209. <https://doi.org/10.1111/gcb.13029>
- 644 Lloret, F., Siscart, D., Dalmases, C., 2004. Canopy recovery after drought dieback in holm-oak
645 Mediterranean forests of Catalonia (NE Spain). *Glob. Chang. Biol.* 10, 2092–2099.
646 <https://doi.org/10.1111/j.1365-2486.2004.00870.x>
- 647 Mohammed, S., Elbeltagi, A., Bashir, B., Alsafadi, K., Alsilibe, F., Alsalmam, A., Zeraatpisheh,
648 M., Széles, A., Harsányi, E., 2022. A comparative analysis of data mining techniques for
649 agricultural and hydrological drought prediction in the eastern Mediterranean. *Comput.*
650 *Electron. Agric.* 197, 106925.
651 <https://doi.org/https://doi.org/10.1016/j.compag.2022.106925>
- 652 Mulero, G., Jiang, D., Bonfil, D.J., Helman, D., 2023. Use of thermal imaging and the
653 photochemical reflectance index (PRI) to detect wheat response to elevated CO₂ and
654 drought. *Plant. Cell Environ.* 46, 76–92. <https://doi.org/10.1111/pce.14472>
- 655 Olano, J.M., García-Cervigón, A.I., Sangüesa-Barreda, G., Rozas, V., Muñoz-Garachana, D.,
656 García-Hidalgo, M., García-Pedrero, Á., 2021. Satellite data and machine learning reveal
657 the incidence of late frost defoliations on Iberian beech forests. *Ecol. Appl.* 31, e02288.
658 <https://doi.org/10.1002/eap.2288>
- 659 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M.,
660 Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: Machine learning in Python.
661 *J. Mach. Learn. Res.* 12, 2825–2830.
- 662 Peñuelas, J., Filella, I., Biel, C., Serrano, L., Savé, R., 1993. The reflectance at the 950–970

- 663 nm region as an indicator of plant water status. *Int. J. Remote Sens.* 14, 1887–1905.
664 <https://doi.org/10.1080/01431169308954010>
- 665 Przeździecki, K., Zawadzki, J.J., Urbaniak, M., Ziemblińska, K., Miatkowski, Z., 2023. Using
666 temporal variability of land surface temperature and normalized vegetation index to
667 estimate soil moisture condition on forest areas by means of remote sensing. *Ecol. Indic.*
668 148, 110088. <https://doi.org/https://doi.org/10.1016/j.ecolind.2023.110088>
- 669 Ritchie, G.A., Hinckley, T.M., 1975. The Pressure Chamber as an Instrument for Ecological
670 Research. *Adv. Ecol. Res.* 9, 165–254. [https://doi.org/10.1016/S0065-2504\(08\)60290-1](https://doi.org/10.1016/S0065-2504(08)60290-1)
- 671 Rog, I., Hilman, B., Fox, H., Yalin, D., Qubaja, R., Klein, T., 2024. Increased belowground
672 tree carbon allocation in a mature mixed forest in a dry versus a wet year. *Glob. Chang.*
673 *Biol.* 30, e17172. <https://doi.org/10.1111/gcb.17172>
- 674 Rog, I., Tague, C., Jakoby, G., Megidish, S., Yaakobi, A., Wagner, Y., Klein, T., 2021.
675 Interspecific Soil Water Partitioning as a Driver of Increased Productivity in a Diverse
676 Mixed Mediterranean Forest. *J. Geophys. Res. Biogeosciences* 126.
- 677 Rouse, J.W., 1973. Monitoring the vernal advancement and retrogradation of natural
678 vegetation. *NASA/GSFC Type II Rep.*
- 679 Sadiq, M.A., Sarkar, S.K., Raisa, S.S., 2023. Meteorological drought assessment in northern
680 Bangladesh: A machine learning-based approach considering remote sensing indices.
681 *Ecol. Indic.* 157, 111233. <https://doi.org/https://doi.org/10.1016/j.ecolind.2023.111233>
- 682 Sanginés de Cárcer, P., Vitasse, Y., Peñuelas, J., Jassey, V.E.J., Buttler, A., Signarbieux, C.,
683 2018. Vapor-pressure deficit and extreme climatic variables limit tree growth. *Glob.*
684 *Chang. Biol.* 24, 1108–1122. <https://doi.org/10.1111/gcb.13973>
- 685 Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least
686 Squares Procedures. *Anal. Chem.* 36, 1627–1639. <https://doi.org/10.1021/ac60214a047>
- 687 Urban, J., Ingwers, M.W., McGuire, M.A., Teskey, R.O., 2017. Increase in leaf temperature
688 opens stomata and decouples net photosynthesis from stomatal conductance in *Pinus taeda*
689 and *Populus deltoides* x *nigra*. *J. Exp. Bot.* 68, 1757–1767.
690 <https://doi.org/10.1093/jxb/erx052>
- 691 Virnodkar, S.S., Pachghare, V.K., Patil, V.C., Jha, S.K., 2020. Remote sensing and machine
692 learning for crop water stress determination in various crops: a critical review. *Precis.*
693 *Agric.* 21, 1121–1155. <https://doi.org/10.1007/s11119-020-09711-9>
- 694 Yuan, W., Zheng, Y., Piao, S., Ciais, P., Lombardozzi, D., Wang, Y., Ryu, Y., Chen, G., Dong,
695 W., Hu, Z., Jain, A.K., Jiang, C., Kato, E., Li, S., Lienert, S., Liu, S., Nabel, J.E.M.S.,
696 Qin, Z., Quine, T., Sitch, S., Smith, W.K., Wang, F., Wu, C., Xiao, Z., Yang, S., 2019.
697 Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Sci. Adv.*
698 5, eaax1396. <https://doi.org/10.1126/sciadv.aax1396>
- 699 Zarco-Tejada, P.J., Camino, C., Beck, P.S.A., Calderon, R., Hornero, A., Hernández-Clemente,
700 R., Kattenborn, T., Montes-Borrego, M., Susca, L., Morelli, M., Gonzalez-Dugo, V.,
701 North, P.R.J., Landa, B.B., Boscia, D., Saponari, M., Navas-Cortes, J.A., 2018. Previsual
702 symptoms of *Xylella fastidiosa* infection revealed in spectral plant-trait alterations. *Nat.*
703 *Plants* 4, 432–439. <https://doi.org/10.1038/s41477-018-0189-7>
- 704 Zarco-Tejada, P.J., Poblete, T., Camino, C., Gonzalez-Dugo, V., Calderon, R., Hornero, A.,
705 Hernandez-Clemente, R., Román-Écija, M., Velasco-Amo, M.P., Landa, B.B., Beck,
706 P.S.A., Saponari, M., Boscia, D., Navas-Cortes, J.A., 2021. Divergent abiotic spectral

- 707 pathways unravel pathogen stress signals across species. *Nat. Commun.* 12, 6088.
708 <https://doi.org/10.1038/s41467-021-26335-3>
- 709 Zhang, L., Jiao, W., Zhang, H., Huang, C., Tong, Q., 2017. Studying drought phenomena in
710 the Continental United States in 2011 and 2012 using various drought indices. *Remote*
711 *Sens. Environ.* 190, 96–106. <https://doi.org/https://doi.org/10.1016/j.rse.2016.12.010>
- 712 Zheng, Y., Wu, S., Xiao, S., Yu, K., Fang, X., Xia, L., Wang, J., Liu, S., Freeman, C., Zou, J.,
713 2022. Global methane and nitrous oxide emissions from inland waters and estuaries. *Glob.*
714 *Chang. Biol.* 28, 4713–4725. <https://doi.org/10.1111/gcb.16233>
- 715

Table 1. Specifications of sensors used in this study: UAV, Sentinel-2, VENμS, and Planet. The temporal and spatial resolutions, along with the specific spectral bands and their corresponding wavelength ranges, are presented. Only the relevant bands to this research are included.

Sensor	Temporal Resolution	Spatial Resolution	Bands	Wavelength Range (nm)
Matrice 600 Pro Hexacopter (UAV)	~1 Month	2–3cm	274 bands (~2 nm bandwidth)	400-1000
Sentinel-2 (satellite)	5 Days	10m	B2 (Coastal Aerosol)	443-490
			B2 (Blue)	490-560
			B3 (Green)	560-665
			B4 (Red)	665-705
		20m	B5 (Vegetation Red Edge)	705-740
			B6 (Vegetation Red Edge)	740-783
			B7 (Vegetation Red Edge)	783-842
		10m	B8 (NIR)	842-865
		20m	B8A (Narrow NIR)	865-945
			B9 (Water vapor)	945-1375
VENμS (satellite)	2 Days	5m	B2	400-440
			B3	423-463
			B4	470-510
			B5	535-575
			B6	600-640
			B7	652-682
			B8	690-714
			B9	734-750
			B10	774-790
			B11	845-885
			B12	900-920
			Coastal Blue	433-453
PlanetScope (satellite)	1 Day	3m	Blue	465-515
			Green I	513-549
			Green	547-583
			Yellow	600-620
			Red	649-680
			Red-Edge	697-712
			NIR	845-885

Table 2. Results of two-sided t-tests comparing the effects of drought and control treatments on leaf water potential (ψ_{leaf}), transpiration (T_r), and carbon assimilation (A_n) for the five woody species across four seasons. P-values from the t-tests are reported, with significant ($p < 0.05$) and marginally significant ($0.1 > p > 0.05$) effects indicated in bold and italics, respectively.

Species	Season	ψ_{leaf} (MPa)	T_r (mmol m $^{-2}$ s $^{-1}$)	A_n (μmol m $^{-2}$ s $^{-1}$)
<i>cypress</i>	Spring	0.771	0.994	0.275
	Summer	<i>0.070</i>	0.910	0.542
	Autumn	0.127	0.366	0.335
	Winter	0.770	0.603	0.160
<i>oak</i>	Spring	0.751	0.127	0.428
	Summer	0.002	0.026	0.101
	Autumn	<i>0.066</i>	0.133	0.472
	Winter	0.666	0.117	0.309
<i>pine</i>	Spring	0.010	0.249	0.927
	Summer	0.247	0.702	0.274
	Autumn	0.589	0.797	0.577
	Winter	0.672	0.904	0.164
<i>carob</i>	Spring	0.314	0.336	0.021
	Summer	0.963	<i>0.089</i>	0.117
	Autumn	0.212	0.472	0.338
	Winter	0.841	0.040	0.643
<i>pistacia</i>	Spring	0.239	0.348	0.754
	Summer	0.978	0.628	0.635
	Autumn	0.381	0.989	0.294
	Winter	0.459	0.454	0.376

Table 3. Performance metrics of the three machine learning models for binary classification using hyperspectral drone imagery models trained using either the full spectral range (274 bands, 400-1000 nm) or a reduced set of the most important features (determined by feature importance ranking, retaining 80% of cumulative importance). Each sample represents the spectral signature of an individual tree canopy. The highest score for each performance metric is highlighted in bold.

Performance Metric	Full Spectral Range (274 Bands)			Reduced Features (80% Importance)		
	LR	SVM	RF	LR (21 Bands)	SVM (13 Bands)	RF (23 Bands)
Accuracy	0.85	0.81	0.59	0.82	0.68	0.62
Recall	0.94	0.85	0.50	0.91	0.82	0.53
Precision	0.78	0.76	0.57	0.76	0.61	0.60
F1 Score	0.85	0.81	0.53	0.83	0.71	0.56

Table 4. Performance metrics of the three machine learning models for binary classification using Satellite synthetic bands for PlanetScope, VENμS, and Sentinel-2. The highest score for each satellite and metric is highlighted in bold.

Satellite	PlanetScope			VENμS			Sentinel-2		
	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF
Accuracy	0.55	0.65	0.61	0.64	0.74	0.54	0.54	0.70	0.51
Precision	0.51	0.63	0.58	0.58	0.70	0.50	0.50	0.66	0.47
Recall	0.68	0.67	0.56	0.74	0.76	0.41	0.62	0.74	0.44
F1 Score	0.58	0.65	0.57	0.65	0.73	0.45	0.55	0.69	0.45

Supplementary Table S1. Optimized hyperparameters for Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF). Models were trained on two datasets: complete hyperspectral data (274 bands) and reduced feature sets comprising the top 80% most important bands. The models classified the samples as control or drought-stressed.

Model	Hyperparameter	274 Bands	80% Importance Bands
LR	C	458	723
	fit_intercept	False	False
	intercept_scaling	0.48	0.55
	max_iter	300	500
	penalty	11	11
	solver	liblinear	liblinear
	tol	8.23e-05	3093e-04
SVM	C	702	1.626
	coef0	3.37	0.23
	degree	2	7
	gamma	0.48	4.61
	kernel	poly	poly
	shrinking	False	False
	tol	2.62e-05	7.3e-05
RF	n_estimators	1400	200
	min_samples_split	5	2
	min_samples_leaf	4	1
	max_features	sqrt	sqrt
	max_depth	30	20
	criterion	entropy	gini

Supplementary Table S2. Optimized hyperparameters for Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF) models implemented using the scikit-learn library in Python. The models were trained on synthetic satellite data derived from the original 274 hyperspectral bands, simulating the spectral bands of three satellite platforms: PlanetScope, VENμS, and Sentinel-2. The models classified the samples as control or drought-stressed.

Model	Hyperparameter	PlanetScope	VENμS	Sentinel-2
LR	C	81.7	3.22	1.70
	fit_intercept	False	True	False
	intercept_scaling	0.823	0.623	1.186
	max_iter	300	200	400
	penalty	l1	none	l2
	solver	liblinear	lbfgs	lbfgs
	tol	1.43e-04	3.09e-05	1.65e-05
SVM	C	702	46.3	46.3
	coef0	3.37	0.75	0.75
	degree	2	2	2
	gamma	0.486	5.58	5.86
	kernel	poly	poly	poly
	shrinking	False	True	True
	tol	2.62e-05	8.17e-05	8.17e-05
RF	n_estimators	1800	200	200
	min_samples_split	2	2	2
	min_samples_leaf	1	1	1
	max_features	sqrt	auto	sqrt
	max_depth	50	50	110
	criterion	gini	gini	entropy

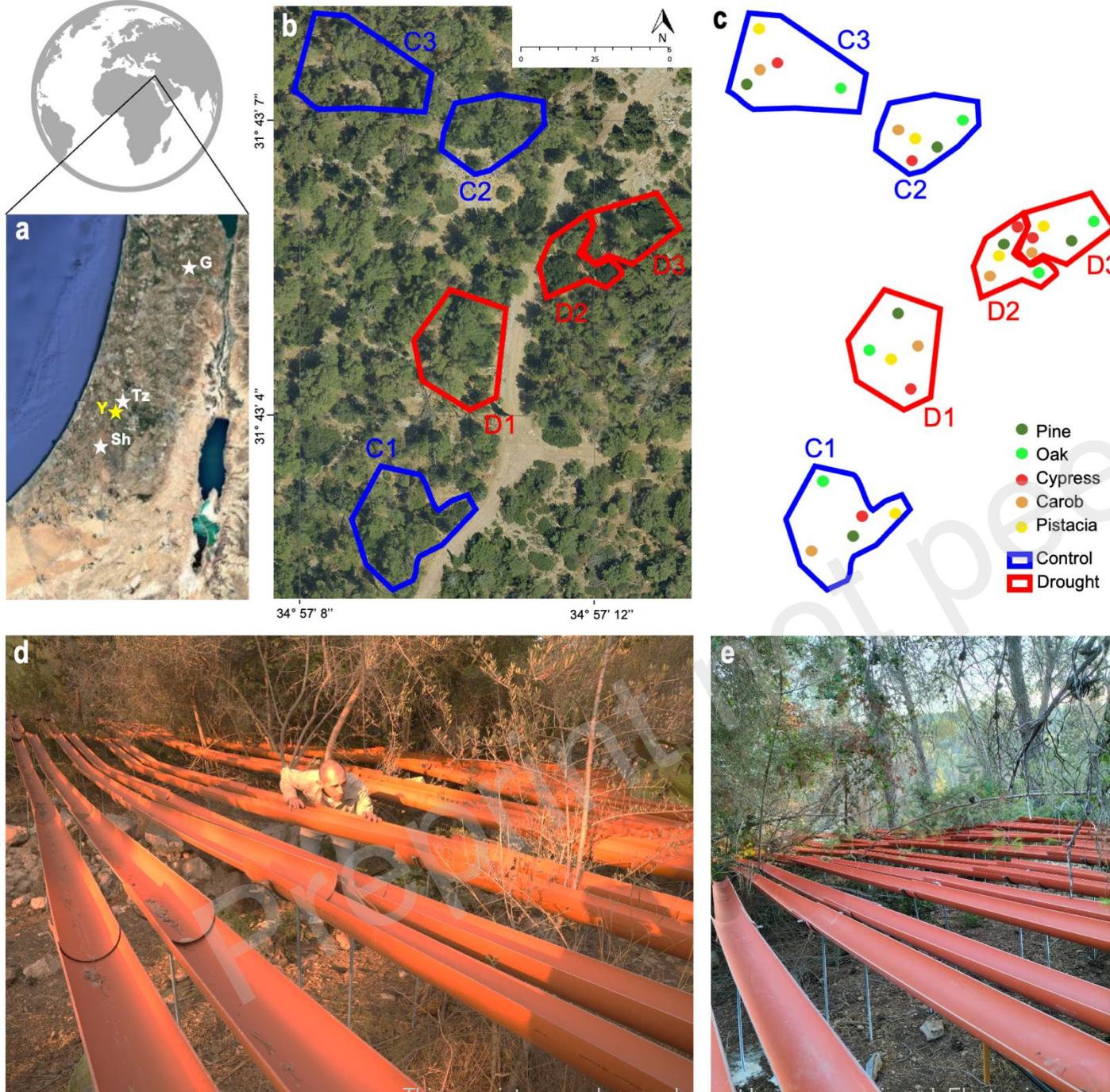


Figure 1. The study area showing (a) the experimental site of Yishi forest (Y; yellow star) and the three forest sites of Tzora (Tz), Shacharia (Sh), and Gilboa (G). (b) Aerial view of the 0.05-hectare plots at Yishi, with control plots in blue (C1-C3) and rainfall exclusion plots in red (D1-D3). (c) Same as (b) but with the tree species marked on the map. (d-e) The rainfall exclusion system in Yishi consisting of open-pipe gutters installed to divert approximately 50% of incident precipitation.

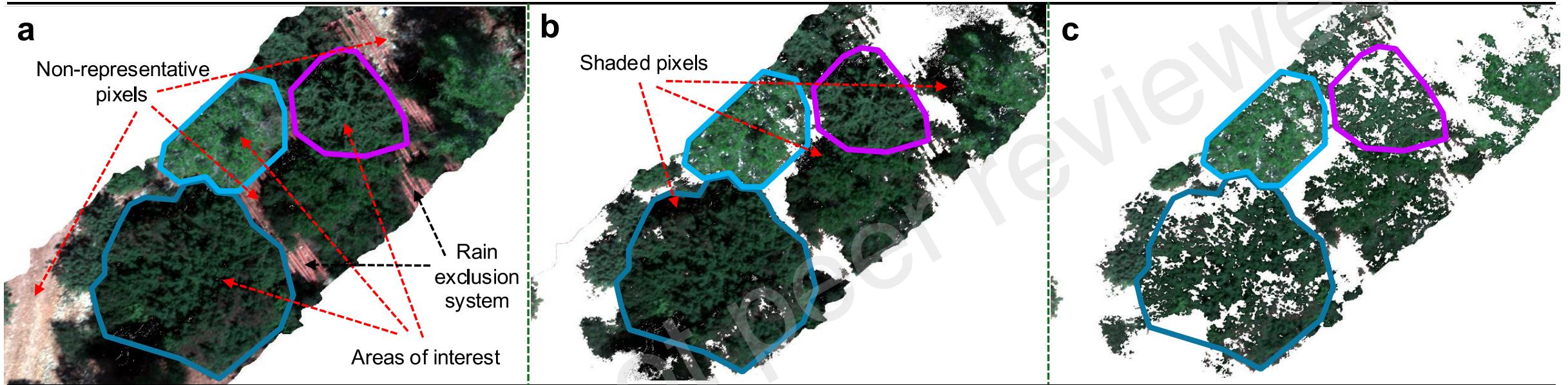
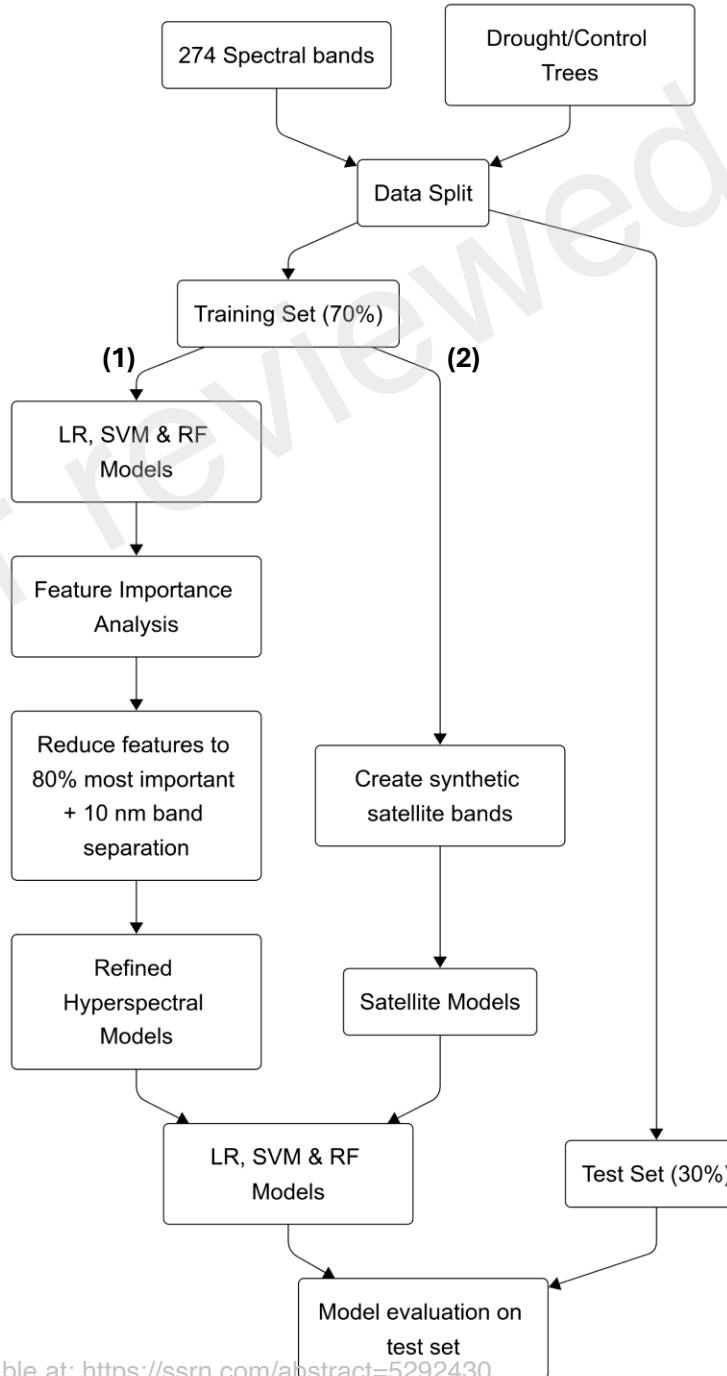


Figure 2. Hyperspectral image preprocessing workflow for isolating tree canopy pixels, showing (a) original image of experimental plots with outlined areas of interest (canopy), non-representative pixels, and visible components of the rainfall exclusion system (black arrows). At first, (b) non-vegetated pixels are removed using an NDVI threshold of <0.3 . (c) Then, to remove the shaded canopy portions, the near-infrared reflectance threshold of $0.07-0.2$ was used.

Figure 3. Workflow of the modeling framework, illustrating the use of inputs comprising 274 spectral bands from hyperspectral imagery and labeled drought/control trees, which were partitioned into training (70%) and test (30%) sets. The training data follows two parallel processing paths: (1) the hyperspectral pathway (left) where machine learning models are trained, followed by feature importance analysis and dimensionality reduction to retain bands representing 80% cumulative importance with 10 nm separation, resulting in refined hyperspectral models; and (2) the satellite simulation pathway (right) where synthetic satellite bands are created to train satellite-compatible models. Both refined hyperspectral and satellite models are evaluated using the same held-out test set to assess classification performance for drought stress detection.



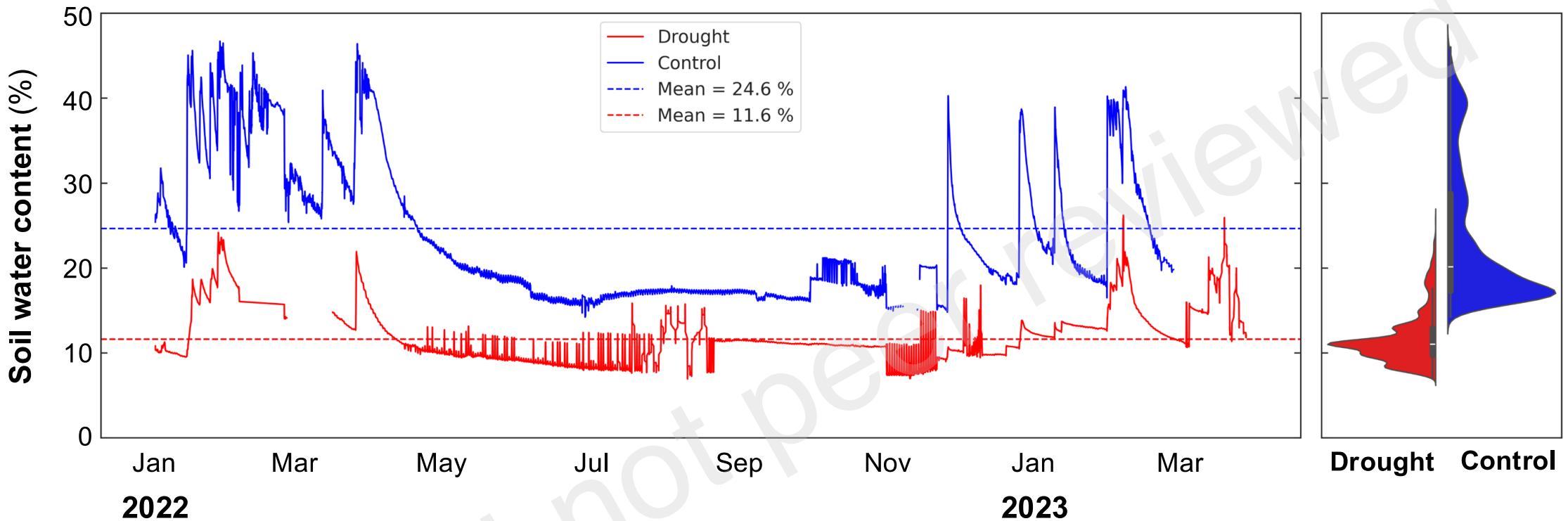


Figure 4. Soil water content (SWC; %) dynamics in control and rainfall exclusion plots over the experimental period. The blue line represents the averaged SWC in control plots, and the red one represents the SWC in rainfall exclusion plots from January 2022 to March 2023. Horizontal dashed lines indicate the mean values across the entire period. The violin plot on the right displays the full distribution of soil moisture measurements for each treatment. Data were smoothed using a Savitzky-Golay filter with a window length of 21 and a polynomial order of 3 to improve the visualization of temporal trends.

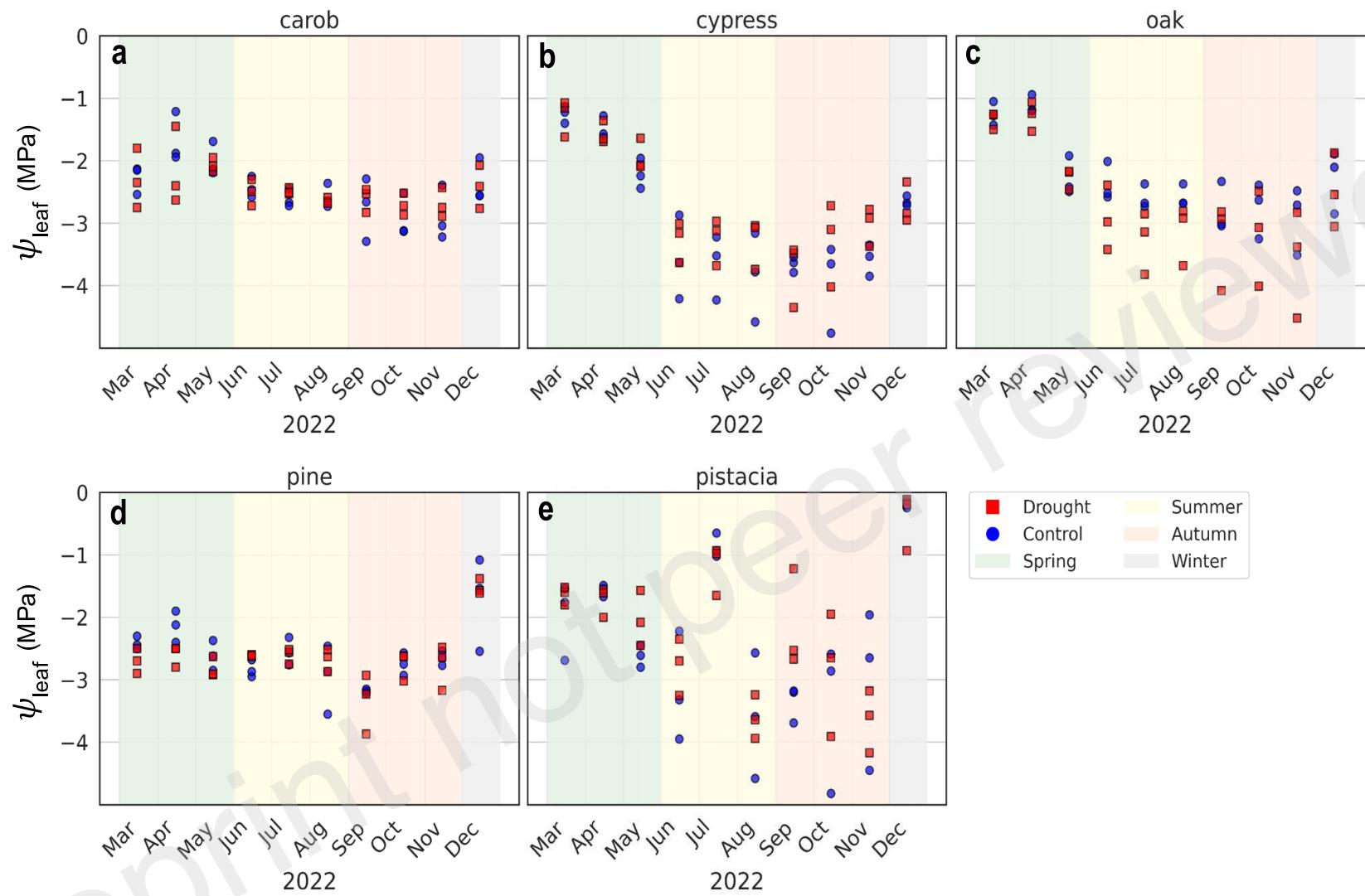


Figure 5. Leaf water potential (ψ_{leaf}) measurements across the five species under control and rainfall exclusion treatments. Scatter plots showing ψ_{leaf} (MPa) for (a) carob, (b) cypress, (c) oak, (d) pine, and (e) pistacia from March to December 2022. Each point represents an individual tree measurement, with control trees denoted by blue circles and drought-treated trees by red squares. Background colors indicate seasons: spring (green, March-May), summer (yellow, June-August), autumn (light red, September-November), and winter (gray, December-February).

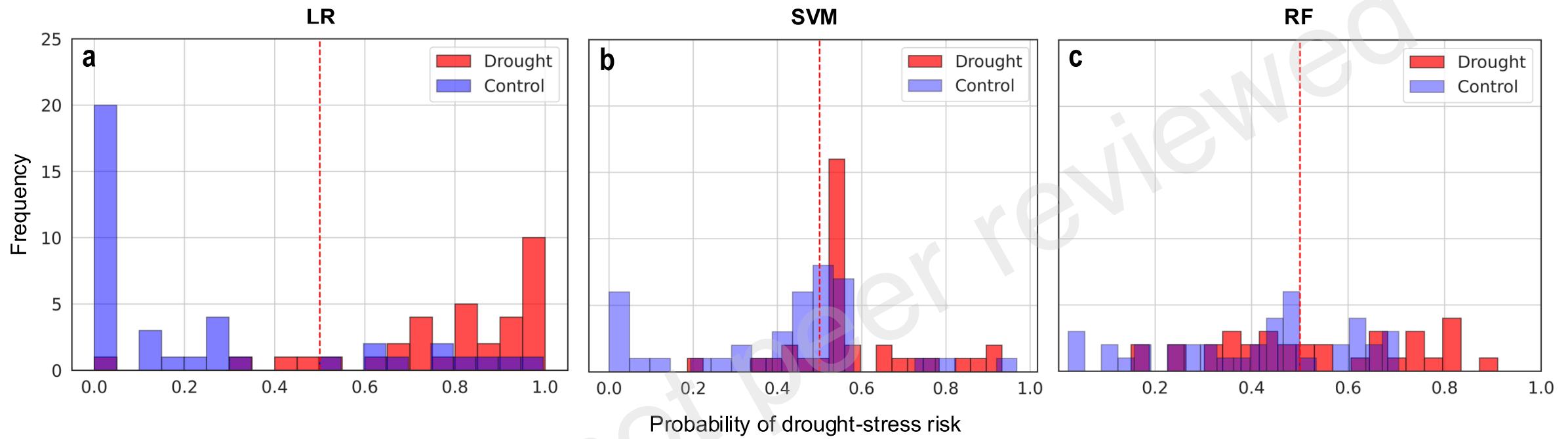
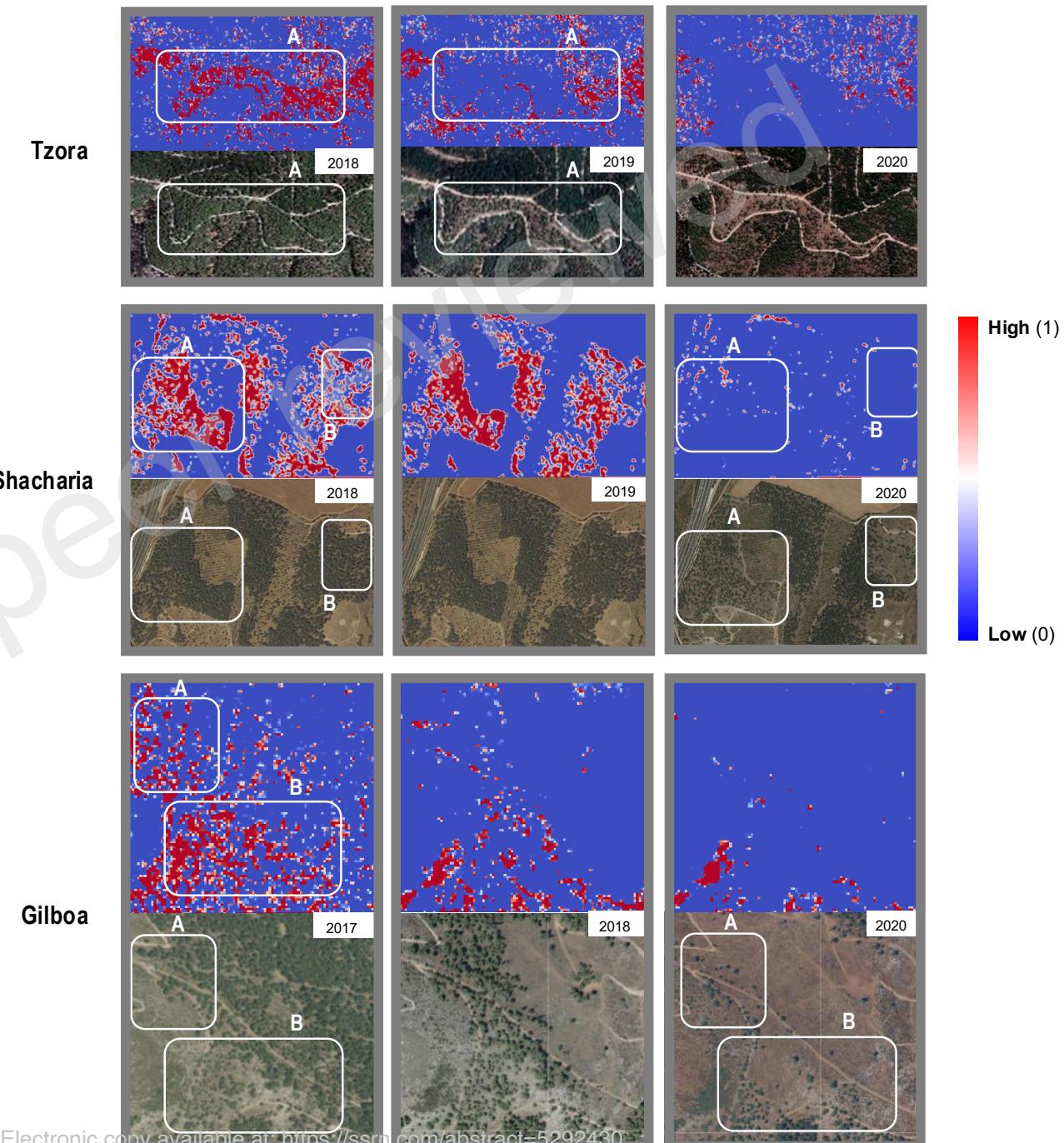


Figure 6. Histograms showing the frequency of predicted drought-stress risk probabilities for **(a)** Logistic Regression (LR), **(b)** Support Vector Machine (SVM), and **(c)** Random Forest (RF) models applied to the test dataset ($N=74$). Models were trained using a reduced hyperspectral feature set that contained only the top spectral bands, representing 80% of the cumulative importance. Blue bars represent control samples and red bars represent drought-treated samples. The vertical dashed red line at 0.5 indicates the classification threshold above which samples were classified as drought-stressed.

Figure 7. Application of SVM drought detection model to VENμS satellite imagery across three Mediterranean forest sites over multiple years. Each row represents a different forest: Tzora (top), Shacharia (middle), and Gilboa (bottom). Each column shows a different year (the year is indicated at the top right corner) with paired images: upper panels display the model's drought risk prediction (blue = low risk, red = high risk) and lower panels show corresponding RGB satellite imagery. White rectangular outlines (A and B) highlight areas of interest where the model predicted high drought stress risk (red pixels) before visible decline was visible. The color bar on the right indicates drought stress risk probability from 0 (low, blue) to 1 (high, red).



Supplement

Preprint not peer reviewed

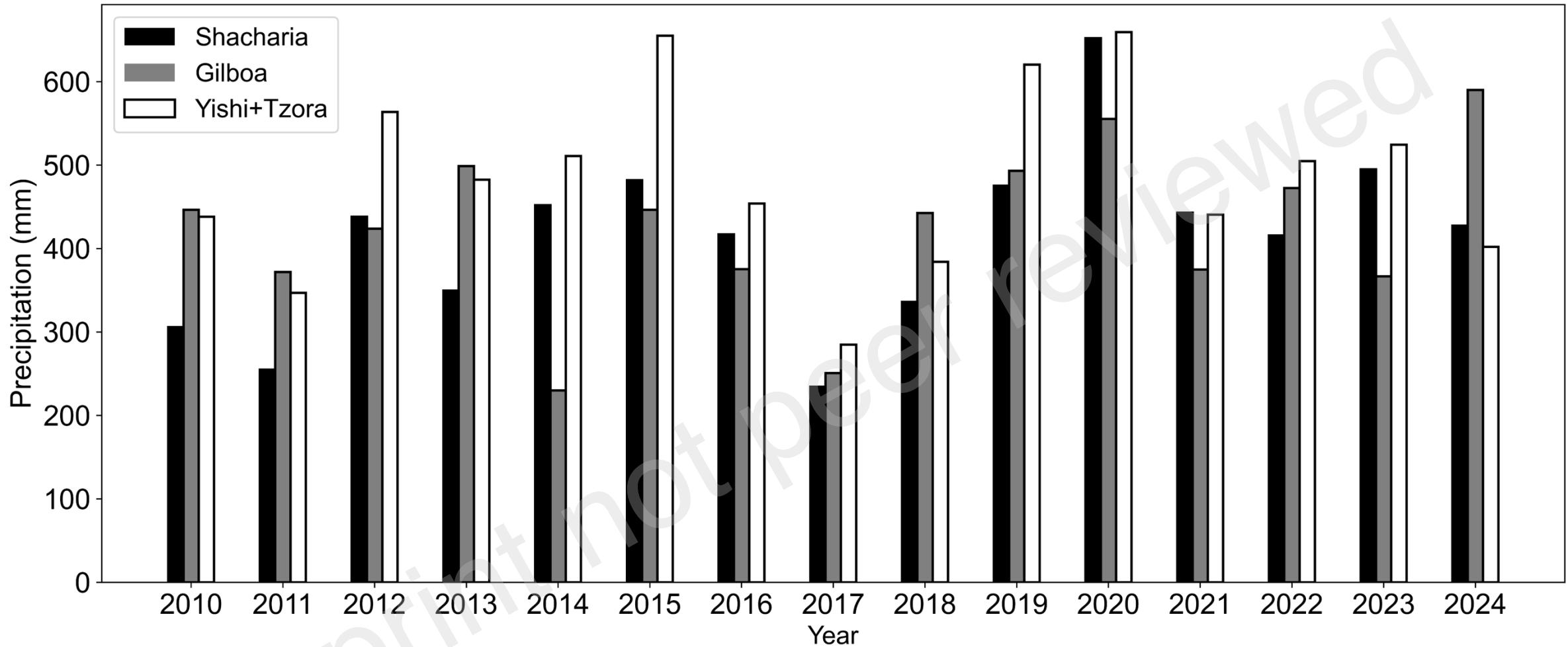


Figure S1. Annual precipitation at the three study sites from 2010 to 2024. Bar chart showing total annual precipitation (mm) recorded at meteorological stations near Shacharia (black), Gilboa (gray), and Yishi+Tzora (white) forests. Data were obtained from the Israel Meteorological Service stations located 4.8 km, 7.5 km, and 5.8 km from the respective forest sites. The year 2017 shows notably lower precipitation across all sites, representing a severe drought event that was used for model validation.

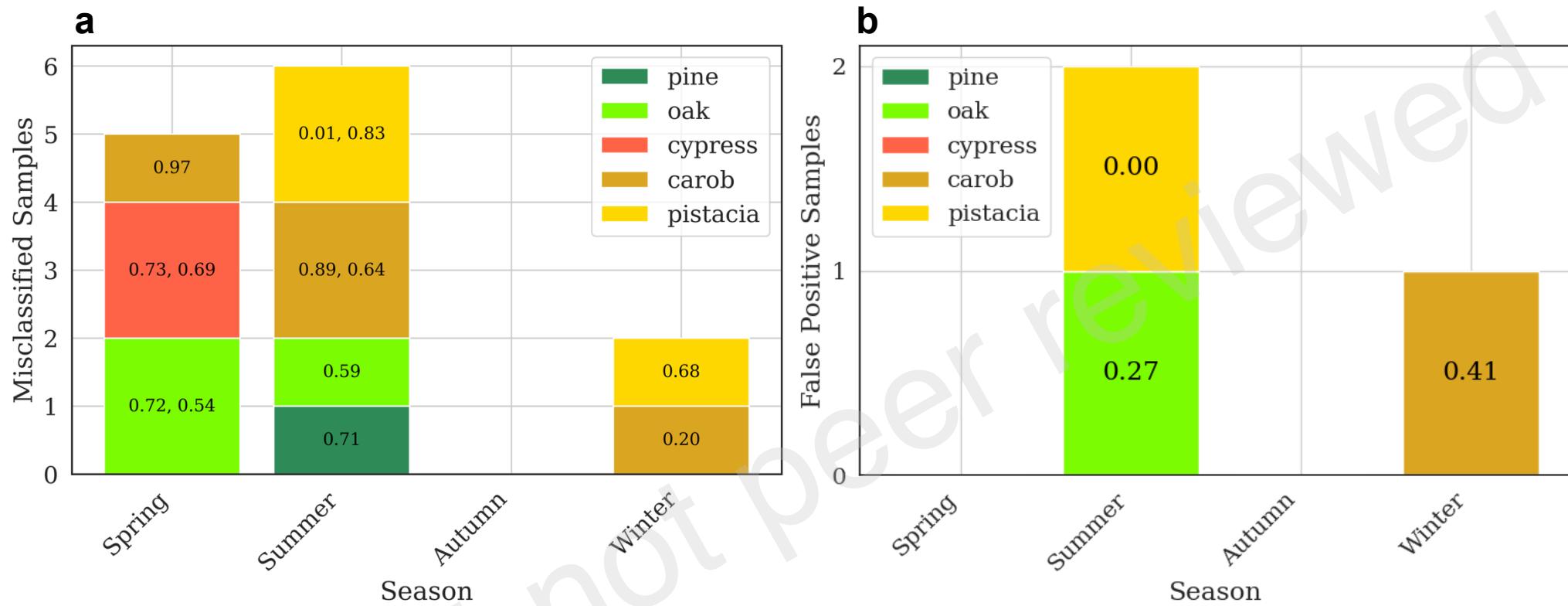


Figure S2. Seasonal distribution of misclassified samples from the Logistic Regression model with 80% feature importance. **(a)** Stacked bar chart showing all misclassified samples by season (spring, summer, autumn, winter) and tree species (pine, oak, cypress, carob, pistacia), with numbers indicating the model's drought probability score for each misclassified sample. **(b)** Stacked bar chart displaying only false positive samples (drought-stressed samples incorrectly classified as control) by season and species, with probability scores shown inside each bar segment. The number represents the score that the model assigned to each sample, where scores above 0.5 (50%) resulted in classification as drought-stressed.

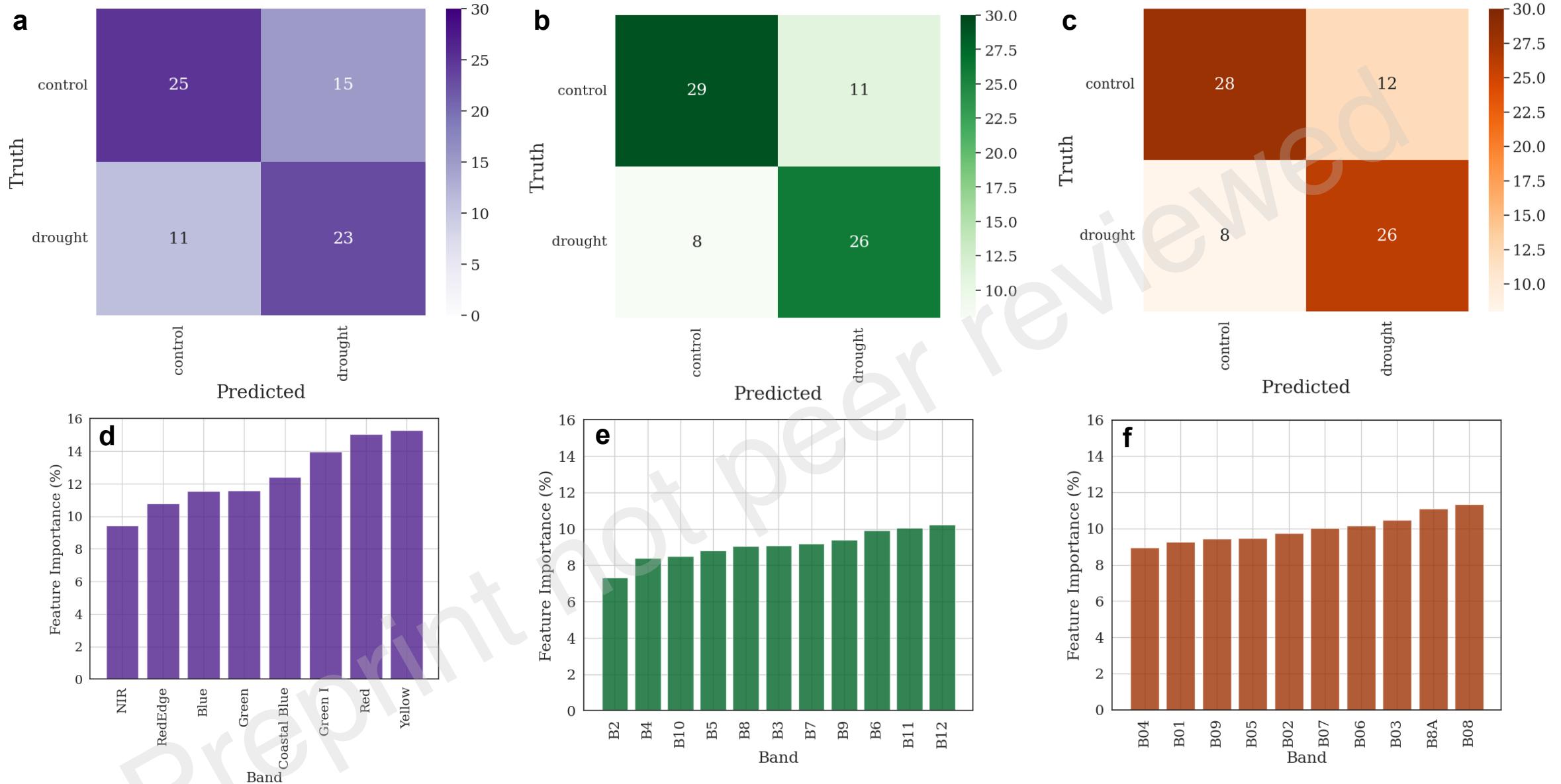


Figure S3. Confusion matrices and feature importance for SVM models trained on synthetic satellite spectral data. The top row shows confusion matrices for (a) Planet, (b) VENμS, and (c) Sentinel-2 satellite bands, displaying the number of samples correctly and incorrectly classified in each category (control vs. drought). Numbers in each cell represent the count of samples. Bottom row shows feature importance (%) for (d) Planet (purple), (e) VENμS (green), and (f) Sentinel-2 (orange) satellite bands.