

סדרת ניירות עבודה
WORKING PAPER SERIES

מס' 100 No.

הערכת טיב המיפוי של שימושי קרקע 2014 בלשכה המרכזית לסטטיסטיקה
Evaluation and Accuracy Assessment of the CBS Land Use Map

אפרת אברהם *

Efrat Avraham *

אלול תשע"ו, ספטמבר 2016 September

***הלשכה המרכזית לסטטיסטיקה – אגף בכיר מערכות מידע**

***Central Bureau of Statistic – Senior Department of Information Technology**

הלשכה המרכזית לסטטיסטיקה (הלמ"ס) מעודדת מחקר המבוסס על נתוני הלמ"ס, כדוגמת עבודה זו. עבודות מחקר אלו אינן פרסומים רשמיים של הלמ"ס, ומכאן שהדעות והמסקנות הבאות בהן לידי ביטוי, הן של המחברים עצמם ואינן משקפות בהכרח את הדעות והמסקנות של הלמ"ס.

הוצאת הלשכה המרכזית לסטטיסטיקה, רח' כנפי נשרים 66, פינת רחוב בקי,

ת"ד 34525, ירושלים 91342

טל': 02-6592666; פקס: 02-65921340

אתר הלמ"ס באינטרנט: www.cbs.gov.il

דואר אלקטרוני: info@cbs.gov.il

תקציר

הלשכה המרכזית לסטטיסטיקה פרסמה נתונים רשמיים חדשים על שימושי הקרקע בישראל. המידע מעודכן לסוף שנת 2013.

בסיס הנתונים עליו מבוסס המידע על שימושי הקרקע הוכן בשנת 2014 באמצעות מערכת המידע הגיאוגרפית (GIS) של הלשכה המרכזית לסטטיסטיקה. העיבוד מבוסס על מידע ממרשמים מנהליים מעוגנים: מרשם דירות ומבנים, מרשם עסקים ומרשם התושבים, שכבות מידע גאוגרפיות שהתקבלו ממשרדים ממשלתיים, מארגונים לא ממשלתיים ומגופים פרטיים ומידע מאורתופוטו לשם טיוב שכבת היערות. מקורות המידע הם: הלשכה המרכזית לסטטיסטיקה, המרכז למיפוי ישראל, הקרן הקיימת לישראל, חברת "מפה", משרד החקלאות ופיתוח הכפר, משרד התשתיות הלאומיות, האנרגיה והמים, משרד החינוך, מגן דוד אדום והרשות הארצית לכבאות והצלה.

בסיס הנתונים נבנה בצורת רשת תאים ארצית רציפה בגודל 100X100 מטרים, ובה הערך של שימוש הקרקע העיקרי בכל תא שטח. דהיינו: שימוש הקרקע שהשטח שלו הוא הגדול ביותר מבין שימושי הקרקע המצויים בתא. קיימים בשכבה 37 שימושי קרקע וכיסויי קרקע. שימושי הקרקע מופו בהתאם לתקן מערך קודים ארצי אחיד למיפוי תכסיות ושימושי קרקע, שהוכן במסגרת הוועדה הבין-משרדית למערכות מידע גאוגרפיות, בהתאם לתקן של האו"ם לסיווג תכסית.

שיטת העבודה: שכבת המבנים הארצית סווגה לפי שימושי קרקע על סמך המרשמים המנהליים המעוגנים ועל סמך מגוון שכבות מידע גאוגרפיות. עבור כל אחד מ-37 שימושי הקרקע וכיסויי הקרקע יוצרה שכבת מידע גאוגרפית על סמך שכבת המבנים המסווגת בתוספת מעטפת ברוחב 20 מטרים ועל סמך שכבות מידע גאוגרפיות נוספות. שכבת היערות טויבה באמצעות מיון מבוקר של אורתופוטו. 37 השכבות הללו שולבו לשכבה אחת לפי סדר היררכי שנקבע בין שימושי הקרקע. הסדר ההיררכי הוא סדר הקודים בשכבת שימושי הקרקע, כאשר קוד מספר אחד הוא במקום הגבוה ביותר בהיררכיה. לאחר איחוד 37 השכבות לשכבה וקטורית אחת היא הותמרה לשכבה רסטורית (0.5 מטרים לפיקסל). נבנתה רשת תאים וקטורית (fishnet) בה גודל התא הוא 100 על 100 מטרים ולכל תא הוענק לפי השכבה הרסטורית הערך של שימוש הקרקע העיקרי המצוי בו. דהיינו: שימוש הקרקע שהשטח שלו הוא הגדול ביותר מבין שימושי הקרקע המצויים בתא. התוצר הסופי כולל שכבה רסטורית של תאים בגודל 100 על 100 מטרים בהם נתון הערך של שימוש הקרקע העיקרי ושכבה וקטורית נקודתית של מרכזי התאים.

המידע על שימושי הקרקע נחוץ לצורך תכנון וקביעת מדיניות במגוון נושאים. כדי שניתן יהיה להסתמך על הנתונים יש לדעת מה מידת המהימנות שלהם.

מטרתה של העבודה הנוכחית להעריך, באמצעות מדדים סטטיסטיים, באיזו מידה קיימת התאמה בין המיפוי בשכבת שימושי קרקע 2014 ובין מידע ממקורות אחרים (אורתופוטו ו-Google Street View) אשר שימשו לאימות תוצאות המיפוי.

יחידת הדגימה, לפיה בוצעה ההשוואה, היתה תא שטח בגודל 100X100 מטרים, כיוון שהמידע על שימושי הקרקע בבסיס הנתונים נתון לפי תאי שטח בגודל זה. בשלב הראשון נלקח מכל קטגוריה מדגם בגודל קבוע, ללא התחשבות בגודל הקטגוריה, כדי לאפשר יכולת ניבוי גם עבור קטגוריות קטנות של שימושי קרקע. בשלב השני, פותחה בעבודה זו שיטה לשקלול של אינדקס Kappa שמאפשר להציג את ה-Kappa, שהיה מתקבל, במונחים של מדגם פרופורציוני. לכל קטגוריה ניתן משקל בהתאם לאחוז שלה מהסה"כ. נבחנו 12 קטגוריות של שימושי קרקע, שכללו את כל שימושי הקרקע בשכבה. נמצא כי הדיוק הכולל של מיפוי הקטגוריות הוא 82.75%. הדיוק הכולל המשוקלל בגודל הקטגוריה הוא 80%. הערך של אינדקס Kappa הוא 0.81 והערך של אינדקס Kappa משוקלל הוא 0.75. עבור הקטגוריות "חפור מופר" ו"גידולי שדה" האמינות של המיפוי היא כ-60% בלבד. עבור שאר הקטגוריות האמינות גבוהה, לרוב היא מעל 85%.

כמו כן, נבחנו 11 שימושי קרקע נבחרים המצויים בשטח הבנוי. נמצא כי הדיוק הכולל של מיפוי שימושי הקרקע הללו הוא 71.52% והדיוק הכולל המשוקלל שלהם הוא 74%. ערך אינדקס Kappa שהתקבל הוא 0.69 והערך של אינדקס Kappa משוקלל הוא 0.7. עבור שימושי הקרקע "שירותי דת", "חירום והצלה", "בריאות ורווחה" ו"שירותי מינהל ציבורי" אמינות המיפוי נמוכה – 60% ומטה. עבור שימושי הקרקע "חינוך והשכלה", "מסחר ומשרדים" ו"תעשייה" אמינות המיפוי בינונית – בין 67% ל-77% ועבור שימושי הקרקע "תרבות פנאי וספורט", "מלונאות ואירוח כפרי", "מסחר ומגורים במרכז עירוני" ו"בתי עלמין" אמינות המיפוי גבוהה – מעל 80%.

פירוט בקשר ליצירת שכבת שימושי הקרקע מצוי בקובץ ההסבר לשכבה (נספח 1). תיאור מפורט יותר של אופן יצירת השכבה, סקירה של שיטות לאיסוף מידע על שימושי קרקע, סקירה של ממצאים בנושא שימושי הקרקע ודוגמאות לשימושים בשכבה יופיעו בעבודה שכתובתה תסתיים בראשית שנת 2017:

אברהם, א' (בהכנה). *יצירת בסיס נתונים ארצי עדכני בנושא שימושי הקרקע באמצעות*

שילוב בין מידע ממרשמים מנהליים ומשכבות מידע גאוגרפיות ובין מידע המופק

באמצעות חישה מרחוק. עבודה למסלול הקידום המקצועי. ירושלים: הלשכה המרכזית

לסטטיסטיקה.

מילות מפתח: שימושי קרקע, כיסוי קרקע, תכסית, אימות, ולידציה.

תודות

ברצוני להודות לד"ר נועם כהן, ראש תחום בכיר באגף בכיר מתודולוגיה סטטיסטית על הייעוץ הסטטיסטי ועל ההדרכה המקצועית, שנתן לי בנוגע לתהליכי הבקרה של בסיס הנתונים המרחבי בנושא שימושי הקרקע ובנוגע למדדים הסטטיסטיים, ששימשו להערכת בסיס הנתונים.

תודתי נתונה לאיל מהריאן, ראש תחום ממ"ג-גאוגרפיה על כך שבחר בי לתכנן את בסיס הנתונים המרחבי בנושא שימושי הקרקע וליצור אותו. כמו כן אני מודה לו על כך שיעץ עצות טובות ותרם מהידע המקצועי שלו.

ברצוני להודות לד"ר רקפת שפרן נתן מאוניברסיטת בן-גוריון בנגב על כך שלימדה אותי ליישם שיטות למיון שימושי קרקע באמצעות נתונים של חישה מרחוק. תודה לפולינה בלכמן מתחום ממ"ג-גאוגרפיה על כך שהצטרפה אלי לשם טיוב שכבת היערות.

תודה לאורן רז מתחום ממ"ג-גאוגרפיה על הפקתה של מפת שימושי קרקע במטרופולין חיפה על סמך בסיס הנתונים שיצרתי.

ברצוני להודות לשחר כ"ץ, מנהל אגף בכיר מערכות מידע על הערותיו והארותיו.

תוכן העניינים

תקציר

1.	מבוא.....	6
2.	מושגים והגדרות.....	7
3.	מקורות המידע ששימשו ליצירת שכבת שימושי קרקע 2014.....	9
4.	שכבת שימושי קרקע 2014 – אופן יצירת השכבה.....	11
5.	סקירת ספרות.....	20
5.1	סוגיות מרכזיות שניתן לבחון בנוגע למיפוי.....	20
5.2	השלבים העיקריים בתהליך אימות נתוני מיפוי.....	20
5.3	שיטות "איכותיות" (qualitative) לאימות נתונים.....	23
6.	מתודולוגיה וממצאים.....	23
6.1	בקורות על מיפוי קטגוריות שימושי הקרקע.....	25
6.1.1	הדיוק הכולל של המיפוי.....	27
6.1.2	אינדקס Kappa.....	28
6.1.3	בדיקת הצלחת המיפוי עבור קטגוריה נתונה של שימוש קרקע ביחס לאימות.....	33
6.2	בקורות על שימושי הקרקע הבנויים שנבחרו.....	37
6.2.1	הדיוק הכולל של המיפוי.....	41
6.2.2	אינדקס Kappa.....	42
6.2.3	בדיקת הצלחת האימות של שימוש קרקע נתון ביחס לאימות.....	44
7.	סיכום ומסקנות.....	47
	רשימת מקורות.....	50
	נספחים.....	54
נספח 1:	קובץ ההסבר לתוצר המדף "שכבת שימושי קרקע 2014"-מס' קטלוגי 10202.....	54
נספח 2:	מפת שימושי קרקע במטרופולין חיפה.....	59
נספח 3:	שימושי הקרקע - קטגוריות לצורך בקרת התוצר.....	60
נספח 4:	שלבי היצירה של נוסחה (7) - אינדקס Kappa משוקלל (Weighted Kappa).....	61

Abstract

1. מבוא

מטרתה של עבודה זו היא להעריך את טיב המיפוי של שימושי הקרקע בבסיס הנתונים הארצי שיוצר בלשכה המרכזית לסטטיסטיקה בשנת 2014.

שימושי קרקע הם הדרכים בהן האדם מנצל את הקרקע: הסדרים, פעילויות ותשומות אנושיות לצורך ניצול הקרקע, שינוי שלה או שימורה.

למידע על שימושי הקרקע חשיבות רבה. המידע נחוץ לצורך תכנון וקביעת מדיניות במגוון נושאים: עיור ובינוי, אספקת שירותים, שמירה על שטחים פתוחים, איכות סביבה, תחבורה, ייעור, חקלאות ועוד.

במטרה להשלים חסרים בסטטיסטיקה הרשמית של מדינת ישראל בנושא קרקע, יוצר בשנת 2014 בתחום ממ"ג-גאוגרפיה בלשכה המרכזית לסטטיסטיקה בסיס נתונים מרחבי בנושא שימושי קרקע¹. המידע מעודכן לסוף שנת 2013.

הפרויקט מספק מידע עדכני ורציף על שימושי הקרקע במדינה² וכן על שימושי הקרקע בעיריות ובמועצות המקומיות ביהודה והשומרון³ ומאפשר הפקת נתונים עבור כל תיחום גאוגרפי⁴, המצוי בשטח שמופה.

בבנתה רשת ארצית רציפה של תאי שטח שגודלם 100 על 100 מטרים. לכל תא הוענק הערך של שימוש הקרקע העיקרי המצוי בו. דהיינו: שימוש הקרקע, שהשטח שלו הוא הגדול ביותר מבין שימושי הקרקע המצויים בתא. בשכבה קיימים 37 שימושי קרקע וכיסויי קרקע שונים שמופו לפי סדר היררכי.

המיפוי התבסס על עיבוד מידע ממרשמים מנהליים (מרשם דירות ומבנים, מרשם עסקים ומרשם התושבים) ומשכבות מידע גאוגרפיות, שהתקבלו ממשרדים ממשלתיים, מארגונים לא ממשלתיים ומגופים פרטיים.

שימושי הקרקע מופו בהתאם למערך קודים ארצי אחיד למיפוי תכסיות ושימושי קרקע (הוועדה הבין-משרדית למערכות מידע גאוגרפיות, 2014), שהוכן במסגרת הוועדה הבין-משרדית למערכות מידע גאוגרפיות, בהתאם לתקן של האו"ם לסיווג תכסית (FAO, 2005). בנספח 2 ניתן לראות דוגמה של מפת שימושי קרקע במטרופולין חיפה, שהופקה לפי המיפוי בשכבת שימושי קרקע 2014. חלק משימושי הקרקע אוחדו לצורך ההצגה במפה. בשכבה קיים כאמור, מידע נפרד עבור 37⁵ שימושי קרקע ותכסיות⁶.

¹ תוצר מדף, מספר קטלוגי 10202. קובץ ההסבר לתוצר מצוי בנספח 1.

² הפרויקט אינו ממפה מחנות צבא ולא שטחים ביטחוניים.

³ בשטחי יהודה והשומרון ניתנים נתונים על שימושי הקרקע רק בתחום השטח הבנוי של עיריות ומועצות מקומיות. לא מופו המועצות האזוריות בשטחי יהודה והשומרון.

⁴ ראו לדוגמה מידע על שימושי הקרקע לפי רשויות מקומיות בקובץ "הרשויות המקומיות 2014", גיליון: "שימושי קרקע 2013, פירוט מלא" http://www.cbs.gov.il/publications16/local_authorities14_1642/excel/p_libud.xls

⁵ רשימת שימושי הקרקע מופיעה בנספח 3.

⁶ תכסית קרקע היא מכלול הפרטים הפיזיים הנראים המצויים על פני השטח.

2. מושגים והגדרות

תכסית קרקע

מכלול הפרטים הפיזיים הנראים הנמצאים על פני השטח.

הערות:

א. קרקע חשופה, סלע חשוף וגופי מים נחשבים גם הם לכיסוי קרקע אף שלכאורה הם אינם תואמים את ההגדרה. זאת כיוון שמדובר בפני השטח עצמו ולא בפרטים כגון צמחייה או שטח בנוי המצויים על פני השטח (FAO, 2005).

ב. הגדרה זו מבוססת על ההגדרה המופיעה בתקן שפרסם האו"ם עבור סיווג של תכסית ועל תקן מערך סמלים ארצי אחיד למיפוי תכסיות ושימושי קרקע שהוכן במסגרת הוועדה הבין-משרדית למערכות מידע גאוגרפיות (הוועדה הבין-משרדית למערכות מידע גאוגרפיות, 2014; FAO, 2005).

שימושי קרקע

שימושי הקרקע הם האופנים שבהם האדם מנצל את הקרקע: הסדרים, פעילויות ותשומות אנושיות, בסוגים שונים של תכסית קרקע, לצורך ניצול הקרקע, שינוי שלה או שימורה. הערות:

א. שימושי הקרקע מציגים את המצב הנוכחי, ואילו ייעודי קרקע – את השימוש העתידי.
ב. דוגמה להבדל בין תכסית קרקע לשימוש קרקע: "שטח בנוי" הנו תכסית קרקע, ואילו "מגורים" או "תעשייה" הם שימושי קרקע.
(הוועדה הבין-משרדית למערכות מידע גאוגרפיות, 2014; FAO, 2005).

חינוך והשכלה – מתחמים המשמשים למטרת חינוך והשכלה: גני ילדים, בתי ספר, אוניברסיטאות, ישיבות וכיו"ב.

בריאות ורווחה – מתחמים המשמשים לבריאות או לרווחה: קופות חולים, בתי חולים, מעונות יום לקשיש, טיפת חלב וכיו"ב.

שירותי דת – מתחמים המשמשים לפולחן דתי: בתי כנסת, כנסיות, מסגדים וכיו"ב.

חירום והצלה – מתחמים המשמשים לשירותי חירום והצלה: תחנות משטרה, תחנות כיבוי אש, תחנות מד"א וכיו"ב.

שירותי מינהל ציבורי – מתחמים המשמשים את גופי השלטון המרכזי והמקומי: משרדי רשויות מקומיות, משרדי ממשלה וכיו"ב.
בתי עלמין – מתחמי בתי עלמין.

תרבות פנאי וספורט – מתחמים המשמשים לתרבות פנאי וספורט: מוזיאונים, בתי קולנוע, ספריות ציבוריות, מתנ"סים, גני חיות, אצטדיונים, מגרשי ספורט, פארקי שעשועים וכיו"ב.

מלונאות ואירוח כפרי – מתחמי מלונות ושירותי לינה.

מסחר ומשרדים – מתחמים המשמשים למסחר ומשרדים.

מסחר ומגורים במרכז עירוני – שטח תחום בו מצויים לפחות 5 מבנים בהם קיים שימוש מעורב של מגורים ומסחר והמרחק בין מבנה למבנה אינו עולה על 50 מטרים. שימוש קרקע זה מצוי בדרך כלל לאורך רחובות ראשיים לדוגמה: רחוב הרצל בראשון לציון או רחוב הרצל ברחובות.

תעשייה – שטחים המשמשים למבנים ומתקנים של תעשייה.

מתחמים בנויים – הקטגוריה כוללת את הקטגוריות ושימושי הקרקע הבאים: מגורים (שטח בנוי למגורי אדם), שירות לציבור (חינוך והשכלה, בריאות ורווחה, שירותי דת, חירום והצלה, שירותי מינהל ציבורי, בית עלמין, תרבות פנאי וספורט, מלונאות ואירוח כפרי, שירות לציבור ללא אפיון – כאשר ידוע שהשימוש של השטח הוא שירות לציבור אך לא ידוע באיזה תחום. לדוגמה: ידוע כי מדובר במבנה ציבורי אך אין מידע האם משמש לחינוך או לבריאות), מסחר ומגורים במרכז עירוני ומסחר ומשרדים.

מכלולי תשתית – הקטגוריה כוללת את הקטגוריות ושימושי הקרקע הבאים: תחבורה (כבישים, חניונים ותחנות מרכזיות, מכלולי תחבורה מסילתית – תחנות רכבת, מכלולי תחבורה ימית – נמלי ים, מכלולי תחבורה אווירית – שדות תעופה, תחנות תדלוק ושירותי דרך), חשמל ותקשורת (לדוגמה מתחמי אנטנות וקליטה), סילוק פסולת (שטח המשמש לאיסוף זבל) ותשתיות רטובות סגורות (מכלי מים מקורים, מכלי גז ונפט).

מבנים חקלאיים – בתי צמיחה, מתחמי גידול בעלי חיים ושטחי אחסון של תוצרת חקלאית.

חפור מופר – שטח חצוב או כרוי בידי אדם. לדוגמה: מחצבה, חפירה ארכיאולוגית.

גופי מים – גופי מים טבעיים (לדוגמה: הכנרת) וגופי מים מלאכותיים כגון: מאגרים פתוחים לאיגום מים, שפכים וקולחין.

גינון לנוי/פארק ציבורי – שטח ציבורי ירוק, מעשה ידי אדם לדוגמה מדשאות.

יער וחורש – שטחים בהם מצוי יער נטוע או חורש טבעי.

מטעים – מטע עצי פרי.

גידולי שדה – שטח של גידול חקלאי שאינו עצים ושטח שעובד ונזנח.

שיחים בתה ועשבוני – שיחים, בתה ועשבוני.

אזור סלעי – פני השטח מכוסים סלע חשוף.

קרקע – חול/דיונות, קרקע חשופה, סחף נחלים פעיל. בשכבה מופו כקרקע גם שטחים שנותרו ללא סיווג לאחר מיפוי שאר שימושי הקרקע וכיסויי הקרקע.

ההגדרות מבוססות על תקן מערך סמלים ארצי אחיד למיפוי תכסיות ושימושי קרקע שהוכן במסגרת הוועדה הבין-משרדית למערכות מידע גאוגרפיות (הוועדה הבין-משרדית למערכות מידע גאוגרפיות, 2014).

בשכבת שימושי קרקע 2014 מופו שימושי קרקע ותכניות. מספר שימושי הקרקע בשכבה גדול ממספר התכניות ולכן יהיו מקומות לאורך העבודה בהם לצורך נוחות ייעשה שימוש במונח "שימושי קרקע" גם עבור תכניות. לדוגמה: "בקרות על קטגוריות שימושי הקרקע".

3. מקורות המידע ששימשו ליצירת שכבת שימושי קרקע 2014

המיפוי התבסס על עיבוד מידע ממרשמים מנהליים מעוגנים⁷ ומשכבות מידע גאוגרפיות, שהתקבלו ממשרדים ממשלתיים, מארגונים לא ממשלתיים ומגופים פרטיים. נעשה שימוש גם באורתופוטו לצורך טיוב שכבת היערות ולצורך מיפוי הרחבות בנייה ביישובים המצויים במועצות אזוריות.

3.1 מרשמים מנהליים מעוגנים

3.1.1 מרשם דירות ומבנים, שנת 2012

מרשם דירות ומבנים שנבנה בלשכה המרכזית לסטטיסטיקה מבוסס על נתוני ארנונה. הוא מכיל רשומות של נכסים המשמשים למגורים ושל נכסים המשמשים למטרות אחרות: חינוך, מסחר, מלונאות וכו'. מרשם דירות ומבנים 2012 מכיל נתונים עבור 202 יישובים. רשומים בו כ-2.1 מיליון נכסים למגורים וכמיליון נכסים שאינם למגורים (סבח, 2013). במיליון סוג נכס מצוי קוד הנכס ותיאור השימוש של הנכס לפי צו הארנונה שמפרסמת כל רשות מקומית. רשומות המרשם מעוגנות לרמת מבנה ואזור סטטיסטי.

3.1.2 מרשם עסקים ארצי, שנת 2012

מרשם עסקים שהוקם בלשכה המרכזית לסטטיסטיקה הוא מאגר רשומות של עסקים, המכיל נתוני העסקה ופדיון חודשיים, בצירוף אינדיקטורים נוספים כגון ענף כלכלי. המרשם ניזון משני קבצים עיקריים: קובץ המתקבל מרשויות מע"מ (מס ערך מוסף) וקובץ המתקבל מהמוסד לביטוח לאומי (הלשכה המרכזית לסטטיסטיקה, 2016). רשומות המרשם מעוגנות לרמת מבנה ואזור סטטיסטי.

3.1.3 מרשם האוכלוסין, 2013

העבודה מבוססת על נתונים מקובצים מתוך מרשם האוכלוסין מאפריל 2013⁸. הקבצה ראשונה: סך הפרטים מתוך האוכלוסייה הרשומים בכל מבנה אליו עוגנו נתונים.

⁷ עיגון הוא קישור רשומה אלפאנומרית לישות גאוגרפית (מרחבית) על סמך מזהים מרחביים. כלומר, שיוך פריט כלשהו למיקומו המדויק בשטח באמצעות המאפיינים שלו. המאפיינים חייבים לכלול נתון אחד לפחות על אודות מיקומו במרחב כגון כתובת המבנה, סמל יישוב, סמל רחוב ו/או מספר בית וכד'.
⁸ הנתונים חושבו לפי מספר הפרטים מבין השייכים לאוכלוסייה. כלומר: שוהים בארץ באופן קבוע.

הקבצה שניה: סך הפרטים מתוך האוכלוסייה הרשומים בכל אזור סטטיסטי אליו עוגנו נתונים.

3.2. שכבות המידע הגאוגרפיות עליהן מבוססת שכבת שימושי קרקע 2014

בעבודה נעשה שימוש במגוון שכבות גאוגרפיות. חלקן יוצרו בלשכה המרכזית לסטטיסטיקה וחלקן התקבלו מארגונים נוספים. בטבלה 1 להלן מצויה רשימת השכבות.

טבלה 1 שכבות המידע הגאוגרפיות עליהן מבוססת שכבת שימושי קרקע 2014

שם הארגון	שם השכבה	שנת עדכון
הלשכה המרכזית לסטטיסטיקה	מבנים ארצית כולל סוג מבנה	2013
הלשכה המרכזית לסטטיסטיקה	רשת תאי שטח בגודל 100 על 100 מטרים	2014
הלשכה המרכזית לסטטיסטיקה	כבישים בין עירוניים	2013
הלשכה המרכזית לסטטיסטיקה	תחנות רכבת	2008
הלשכה המרכזית לסטטיסטיקה	נקודות ציון מרכזיות ביישובים ערביים	2008
הלשכה המרכזית לסטטיסטיקה	מוקדי תעסוקה וחינוך	2002
הלשכה המרכזית לסטטיסטיקה	מסחר ומגורים במרכז עירוני - עיבוד על סמך מבנים המוגדרים "מגורים בבנין מעורב" בשכבה הנקודתית "מוסדות ושימושי מבנה" של המרכז למיפוי ישראל	2012
הלשכה המרכזית לסטטיסטיקה	תאי דגימה	2008
הלשכה המרכזית לסטטיסטיקה	אזורים סטטיסטיים 2011	2011
הלשכה המרכזית לסטטיסטיקה	מרקמים עירוניים	2010
המרכז למיפוי ישראל	מבנים ארצית	2012
המרכז למיפוי ישראל	מוסדות ושימושי מבנה	2012
המרכז למיפוי ישראל	מכלולים בנויים	2012
המרכז למיפוי ישראל	מתחמים בנויים	2012
המרכז למיפוי ישראל	מכלולי תחבורה	2012
המרכז למיפוי ישראל	מכלולי תשתית	2012
המרכז למיפוי ישראל	אתרים	2012
המרכז למיפוי ישראל	גופי מים	2012
המרכז למיפוי ישראל	כיסוי קרקע	2006

2013	חלקות קדסטר	המרכז למיפוי ישראל
2013	יערות	הקרן הקיימת לישראל
2013	חורש	הקרן הקיימת לישראל
2008	תחנות כיבוי אש	הרשות הארצית לכבאות והצלה
2013	מבנים ציבוריים ומוסדות	חברת "מפה"
2013	חניונים	חברת "מפה"
2013	תחנות דלק	חברת "מפה"
2013	בתי עלמין	חברת "מפה"
2013	מוקדי עניין	חברת "מפה"
2013	מתחמים בנויים	חברת "מפה"
2008	תחנות מגן דוד אדום (מד"א)	מגן דוד אדום
2012	בתי ספר	משרד החינוך
2012	חלקות צומח	משרד החקלאות ופיתוח הכפר
2012	מבני משק לגידול בעלי חיים	משרד החקלאות ופיתוח הכפר
2013	מחצבות פעילות	משרד התשתיות הלאומיות, האנרגיה והמים

3.3 גיליונות אורתופוטו⁹

גיליונות אורתופוטו ברזולוציה מרחבית של 0.5 על 0.5 מטרים לפיקסל משנת 2012 שימשו לטיוב שכבת היערות וגיליונות אורתופוטו ברזולוציה מרחבית של 0.5 על 0.5 מטרים לפיקסל משנת 2014 שימשו להוספת הרחבות בנייה ביישובים המצויים במועצות אזוריות.

4. שכבת שימושי קרקע 2014 – אופן יצירת השכבה

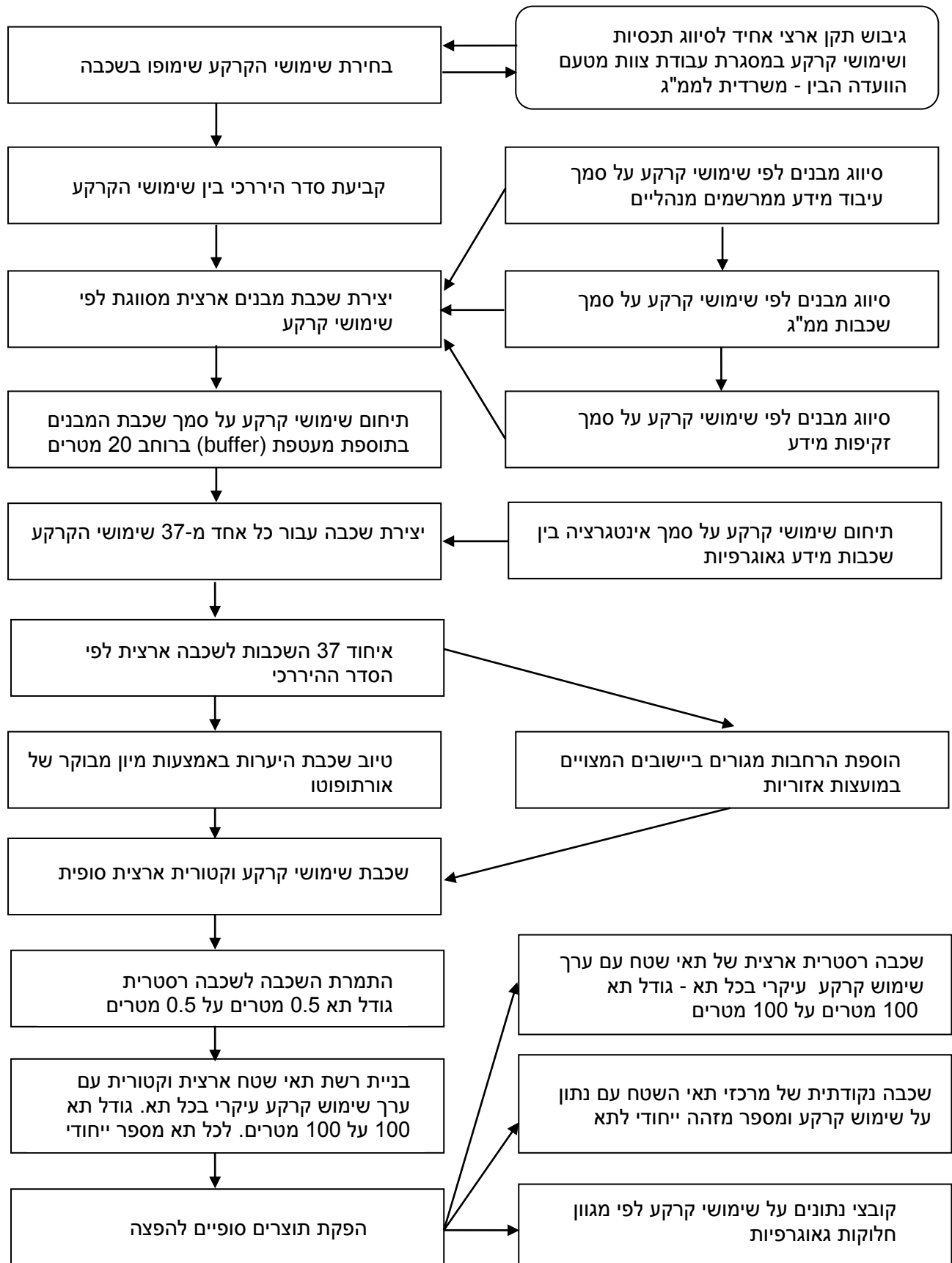
תיאור מפורט של אופן יצירת שכבת שימושי קרקע 2014 יופיע בעבודה אשר כתיבתה תסתיים בראשית שנת 2017 (אברהם, בהכנה)¹⁰. בפרק זה יינתן תיאור מקוצר של השיטה. באיור 1 מופיע תרשים של השלבים העיקריים בתהליך יצירת שכבת שימושי קרקע 2014.

שלב מקדים: גיבוש תקן ארצי אחיד לסיווג תכסיות ושימושי קרקע

שימושי הקרקע המופיעים בשכבת שימושי קרקע 2014, מופו בהתאם לתקן מערך קודים ארצי אחיד למיפוי תכסיות ושימושי קרקע (הוועדה הבין-משרדית למערכות מידע גאוגרפיות, 2014).

⁹ אורתופוטו הוא צילום אוויר סרוק ממנו סולקו העיוותים הגאומטריים שנגרמו על ידי הטופוגרפיה והוא עוגן לרשת קואורדינטות (לוי, 2003). המקור לאורתופוטו הוא המרכז למיפוי ישראל.
¹⁰ העבודה תכלול גם סקירה על שיטות איסוף מידע על שימושי קרקע, סקירה של ממצאים על שימושי הקרקע, דוגמאות לשימושים בשכבה ועוד.

איור 1: השלבים העיקריים בתהליך יצירת שכבת שימושי קרקע 2014



תקן זה הוכן על ידי צוות מטעם הוועדה הבין משרדית למערכות מידע גאוגרפיות, בהתאם לתקן של האו"ם לסיווג תכסית (FAO, 2005). הוא נועד ליצור מינוח אחיד עבור יצרני המידע ועבור המשתמשים במידע בנושא שימושי הקרקע. התקן אמנם אינו משמש מדריך להקמת שכבה ממ"גית מסוימת אולם גיבוש התקן מהווה שלב מקדים ליצירת שכבת שימושי הקרקע.

קיימים קשרים הדדיים בין יצירת השכבה ובין יצירת התקן. מצד אחד, שימושי הקרקע המופיעים בשכבה נקבעו בהתאם לסיווגים המופיעים בתקן. מצד שני, במהלך יצירת התקן השפיעו נציגי הלשכה המרכזית לסטטיסטיקה בצוות שגיבש את התקן על קביעת שימושי הקרקע שיופיעו בתקן. כמו כן, על סמך בדיקת היתכנות, שנערכה בלשכה המרכזית לסטטיסטיקה בקשר למיפוי של שימושי קרקע ותכסיות שהופיעו בתקן, שונו חלק מהסיווגים המצויים בו. שלב זה מופיע באיור 1 בתוך מלבן מעוגל כיוון שהוא שלב מקדים ליצירת השכבה וכן כיוון שהתקן גובש על ידי צוות מטעם הוועדה הבין משרדית למערכות מידע גאוגרפיות.

4.1 בחירת שימושי הקרקע שימופו בשכבה

בתקן הוגדרו שש רמות פירוט עבור נתונים בנושא תכסיות ושימושי קרקע. במסמך שפרסמה הוועדה הבין-משרדית (2014), מצוי התקן וכן הסברים בקשר לקטגוריות ולרמות הפירוט. ההחלטה עבור אלו שימושי קרקע או תכסיות לספק מידע ולפי איזו רמת פירוט נתונה בידי כל אחד מיצרני המידע. בבניית שכבת שימושי קרקע הושם דגש על מיפוי שימושי קרקע בנויים ועל מיפוי שימושי קרקע עיקריים בשטח הפתוח. רבים משימושי הקרקע בקטגוריה "מופר מנוהל" מופו בשכבת שימושי הקרקע לפי רמת הפירוט הששית, הרמה הגבוהה ביותר. לדוגמה "חינוך והשכלה", "תרבות פנאי וספורט" ויתר שימושי הקרקע המצויים תחת הסיווג "שירות לציבור" הנם דוגמה לשימושים שמופו לפי רמת הפירוט השישית. שימושי הקרקע¹¹ בשטח הפתוח מופו בשכבה לפי רמת הפירוט הרביעית או לפי רמת הפירוט החמישית של התקן. כך לדוגמה מופו "יער וחורש" ו-"גופי מים" לפי רמת הפירוט הרביעית. הבחירה ברמת הפירוט, לפיה מופה כל אחד משימושי הקרקע הושפעה גם מרמת הפירוט בה היו נתונים מקורות המידע שנאספו.

4.2 קביעת סדר היררכי בין שימושי הקרקע

נקבעה היררכיה בין שימושי הקרקע. ההיררכיה נועדה לקבוע מהו שימוש הקרקע שימופה במקרים בהם קיים קונפליקט בין מספר שימושי קרקע. קונפליקט בין שימושי קרקע יכול להיווצר כאשר בשטח נתון קיים יותר משימוש קרקע אחד. לדוגמה, כאשר מצוי קולנוע

¹¹ כפי שנאמר בפרק של ההגדרות יהיו מקומות לאורך העבודה בהם לצורך נוחות ייעשה שימוש במונח "שימושי קרקע" גם עבור תכסיות.

(שימוש קרקע "תרבות פנאי וספורט") באזור תעשייה (שימוש קרקע "תעשייה") נוצר קונפליקט בין שני שימושי קרקע. קונפליקט עלול להיווצר גם בין מקורות מידע שונים. לדוגמה, לפי מקור מידע אחד מבנה משמש לחינוך ולפי מקור מידע אחר הוא משמש לתרבות פנאי וספורט. בנספח 1 מצויה טבלה בה רשומים שימושי הקרקע לפי הסדר ההיררכי שנקבע. קוד מספר "1" מייצג את המקום הגבוה ביותר בהיררכיה¹². ניתן לראות כי שימושי קרקע בשטח בנוי קודמים בהיררכיה לשימושי קרקע בשטח פתוח. הסיבה היא שבדרך כלל שימושי הקרקע בשטח בנוי מעודכנים יותר ותחומים בצורה מדויקת יותר בהשוואה לשימושי קרקע בשטח פתוח.

4.3 סיווג שכבת המבנים הארצית לפי שימושי קרקע

המבנים היו יחידת ההכללה הקטנה ביותר עליה מבוססת שכבת שימושי הקרקע. במקורות המידע קיימים נתונים ברמת פירוט גדולה יותר מזו של מבנה. מספר נכסים במבנה נתון יופיעו כרשומות נפרדות במרשם דירות ומבנים, מספר עסקים המצויים באותו מבנה יופיעו כרשומות נפרדות במרשם עסקים ותושבים הגרים במבנה נתון יופיעו במרשם האוכלוסין ברשומות נפרדות.

שכבת המבנים הארצית סווגה לפי שימוש הקרקע העיקרי של המבנה. לאחר מכן, נתחם השטח של שימוש קרקע נתון, באמצעות יצירת buffer¹³ סביב מבנים, המסווגים לפי שימוש קרקע זה. לשטח, שתוחם באמצעות ה-buffer, ניתן הערך של שימוש הקרקע, לפיו סווגו המבנים.

שיטה זו, המבוססת על סיווג המבנים ועל תיחום שימושי הקרקע לפי מרחק קבוע סביב מבנים שסווגו לפי אותו שימוש קרקע נועדה להבטיח אחידות במיפוי שימושי הקרקע. תיחום שימוש קרקע נתון לפי שיטה זו נקבע על ידי גודל המבנים סביבם מיוצר ה-buffer וכן לפי הצורה בה מפוזרים המבנים על פני המרחב.

תהליך סיווג המבנים לפי שימושי קרקע התבצע בשלושה שלבים. בשלב הראשון סווגו המבנים בשכבת המבנים הארצית על סמך עיבוד מידע ממרשמים מנהליים. בשלב השני סווגו מבנים משכבת המבנים הארצית, על סמך מידע משכבות מידע גאוגרפיות. מבנים שלא סווגו לאחר ביצוע שני השלבים הראשונים, סווגו בשלב השלישי לפי שימושי קרקע על סמך זקיפות מידע.

סיווג מבנים לפי שימושי קרקע על סמך עיבוד מידע ממרשמים מנהליים:

סיווג המבנים לפי מרשם דירות ומבנים: המילון הארצי של סוגי הנכסים הוא חיבור בין מילון סוג נכס של הרשויות המקומיות. הקודים שמקצה כל רשות מקומית לתיאור נכס וכן תיאורי

¹² בטבלה זו מופיע עבור כל שימוש קרקע הקוד שניתן לו בתקן.

¹³ Buffer הוא שטח שמוגדר על ידי מרחק נתון סביב יישות גאוגרפית. גודל ה- buffer 20 מטרים.

הנכסים שונים בין הרשויות השונות. לפיכך קודדו כל סוגי הנכסים המופיעים במילון הארצי לפי הקודים של שימושי הקרקע.

קביעת שימוש הקרקע העיקרי לפי מרשם דירות ומבנים נעשתה בשני שלבים: בשלב הראשון נערכה אבחנה בין מבנים למגורים למבנים שאינם למגורים. כאשר במבנה נתון מספר הרשומות שסווגו כמגורים היה גדול או שווה למספר הרשומות שאינן למגורים סווג המבנה כ"מגורים"¹⁴. בשאר המקרים סווג המבנה כ"לא למגורים". בשלב השני הוגדר שימוש קרקע עיקרי עבור המבנים שסווגו בשלב הראשון כ"לא למגורים". השימוש הוגדר לפי שימוש הקרקע שמספר הנכסים שסווגו לפיו הוא הגדול ביותר¹⁵. במקרים בהם היו במבנה נתון מספר זהה של רשומות משני שימושי קרקע הוגדר שימוש הקרקע לפי הסדר ההיררכי שנקבע בין שימושי הקרקע.

סיווג המבנים לפי מרשם עסקים: העסקים המופיעים במרשם העסקים מסווגים לפי ענף כלכלי. בשלב הראשון סווגו הרשומות של העסקים במרשם העסקים לפי שימושי קרקע על סמך השיוך שלהן לענף כלכלי. ישנם מבנים בהם מצויים עסקים אך השימוש העיקרי שלהם הוא "מגורים"¹⁶ לכן בשלב השני נבנה אלגוריתם הקובע עבור כל מבנה אם הוא ימופה כמבנה מגורים או כמבנה המשמש לעסקים. כאשר השימוש לעסקים, נקבע סוג השימוש העיקרי (מסחר, חינוך, תעשייה וכו'). האלגוריתם משקלל את מספר העסקים מכל סוג של שימוש קרקע, המצויים בכל מבנה וכן את מספר התושבים במבנה, לפי מרשם האוכלוסין וכן היררכיה בין שימושי הקרקע. פרטים על התהליך יופיעו אצל אברהם (בהכנה).

בסוף התהליך סווג כל מבנה לפי המידע במרשמים המנהליים וכן לפי המידע בשכבות המידע הגאוגרפיות. זאת לפי סוג המבנה (בית ספר, מוזיאון, קופת חולים) או לפי שכבות מידע שאינן מייצגות מבנים (למשל מבנה המצוי בתוך מתחם של אזור תעשייה סווג לפי שימוש הקרקע "תעשייה").

לאחר שסווגו המבנים על סמך המידע המצוי במרשמים המנהליים ועל סמך המידע המצוי בשכבות המידע הגאוגרפיות היו מבנים שנותרו ללא סיווג. הם סווגו על סמך זקיפות מידע. לדוגמה מבנים המצויים ביישובים ערביים בחלקה בה עוגן מבנה למגורים ולא ניתן להם סיווג לפי שימוש קרקע אחר סווגו כ"מגורים".

4.4 יצירת שכבת מידע גאוגרפית עבור כל אחד משימושי הקרקע

יוצרו 37 שכבות. שכבה עבור כל אחד משימושי הקרקע. מבנים שסווגו לפי שימוש קרקע נתון הוקפו במעטפת (buffer) ברוב 20 מטרים. לפוליון שנוצר ניתן הערך של שימוש הקרקע בהתאם לשימוש הקרקע של המבנים המצויים בו. אם למשל היו מספר מבנים

¹⁴ לדוגמה מבנה שיש בו 8 דירות מגורים ובנק יוגדר כ"מגורים".

¹⁵ לדוגמה שימוש הקרקע העיקרי של קניון הוא מסחר למרות שעשויים להיות בו גם נכסים מסוג אחר כמו למשל מרפאות.

¹⁶ לדוגמה חנות המצויה במבנה מגורים.

המשמשים למסחר שימוש הקרקע של הפוליון שנוצר סביבם לאחר הוספת ה-buffer מופה כמסחר. הוספת ה-buffer נועדה לכלול גם שטחים המצויים בין מבנים. כביש שעובר בתוך אזור תעשייה לדוגמה סווג גם הוא כתעשייה.

נוסף לתיחום באמצעות מבנים בתוספת ה-buffer נעשה שימוש גם בשכבות נוספות. לדוגמה: כדי למפות את שימוש הקרקע "תרבות פנאי וספורט" יש להוסיף לא רק שטחים סביב מבנים המשמשים לתרבות אלא גם שטח של גני חיות, אצטדיונים וכו'. בשלב זה הוכנו גם שכבות מידע עבור שטחים פתוחים לדוגמה שכבה עבור "מטעים".

4.5 איחוד 37 השכבות לשכבה אחת של שימושי קרקע

שכבות המידע אוחדו לשכבה אחת לפי הסדר ההיררכי¹⁷ שנקבע בין השכבות.

4.6 טיוב שכבת היערות באמצעות מיון מבוקר של אורתופוטו

בשכבת היערות של קק"ל מצויים גם שטחים שהתכסית שלהם היא "קרקע". בשטחים הללו לא נראים עצים באורתופוטו או שיש עצים בודדים. לכן לצורך המיפוי בשכבת שימושי קרקע נערך בתחום של שכבת היערות פיענוח של אורתופוטו¹⁸ באמצעות מיון מבוקר (supervised classification) לפי הערכים של ערוצי ה-RGB בגליונות האורתופוטו. המיון נערך בשיטת Maximum Likelihood Classification ונפו משכבת היערות תאי שטח בגודל 100 על 100 מטרים בהם השטח שפוענח כ"עצים" לא היה הגדול ביותר מבין כיוויי הקרקע שפוענחו בתא השטח. במילים אחרות: מתוך שכבת יער קק"ל מופו בשכבת שימושי קרקע כיער רק תאי שטח בגודל 100 על 100 מטרים בהם השטח שפוענח כ"עצים" היה הגדול ביותר מבין כיוויי הקרקע שפוענחו בתא השטח (majority).

4.7 הוספת הרחבות מגורים ביישובים המצויים במועצות אזוריות

שכבת המבנים הארצית עליה מבוססת שכבת שימושי הקרקע מעודכנת לסוף שנת 2013 (ינואר 2014). עיקר פעולת העדכון מתבצעת ביישובים הגדולים – עיריות ומועצות מקומיות. לכן כדי שגם הרחבות מגורים ביישובים המצויים במועצות האזוריות ימופו בשכבת שימושי הקרקע אותרו ההרחבות הללו והוספו על ידי דיגיטציה¹⁹ מאורתופוטו²⁰.

¹⁷ כאמור, בנספח 1 מצויה טבלה בה רשומים שימושי הקרקע לפי הסדר ההיררכי שנקבע. קוד מספר "1" מייצג את המקום הגבוה ביותר בהיררכיה.

¹⁸ הפיענוח באמצעות מיון מבוקר בוצע לפי אורתופוטו משנת 2012 ברזולוציה של 0.5 על 0.5 מטרים לפיקסל.

¹⁹ תהליך בו נקודות, קווים או פוליגונים (מצולעים) נקלטים ממפת נייר (אנלוגית) ומועברים לפורמט דיגיטלי. ניתן לבצע דיגיטציה ישירות ממפת נייר על גבי לוח דיגיטציה או על גבי המסך. כאשר מבצעים דיגיטציה על גבי מסך המחשב, מוצגת ברקע מפה סרוקה, צילום אוויר או הדמאת לוויין (לויין, 2003).

²⁰ ההרחבות הוספו לפי אורתופוטו משנת 2014 ברזולוציה של 0.5 על 0.5 מטרים לפיקסל.

4.8 סיום תהליך יצירת שכבת שימושי הקרקע

לאחר שהוספו לשכבת שימושי הקרקע הארצית (סעיף 4.5) הרחבות הבנייה ביישובים במועצות האזוריות (סעיף 4.7) ובוצעה פעולה של "טיוב" שכבת היערות (סעיף 4.6) יוצרה שכבת שימושי קרקע וקטורית²¹ ארצית סופית, דוגמה באיור 2. שכבה זו הותמרה לשכבה רסטורית²² בה גודל תא הנו 0.5 מטרים על 0.5 מטרים, דוגמה באיור 3. נבנתה רשת תאי שטח ארצית וקטורית (fishnet) בה גודלו של תא שטח הוא 100 על 100 מטרים. לכל תא ניתן מספר תא ייחודי והוענק לו ערך של שימוש הקרקע העיקרי בתא. ערך זה הוענק לו מהשכבה הרסטורית שיוצרה בשלב הקודם (גודל תא 0.5 על 0.5 מטרים). הוא ניתן לפי ה-majority של שימוש הקרקע שמספר הפיקסלים בגודל 0.5 על 0.5 מטרים הוא הגדול ביותר מבין שימושי הקרקע המצויים בתא, דוגמה באיור 4. המשמעות היא שניתן לתא השטח הערך של שימוש הקרקע שהשטח שלו הוא הגדול ביותר מבין שימושי הקרקע המצויים בתא. בשלב האחרון הותמרה רשת תאי השטח של 100 על 100 מטרים עם ערך שימוש הקרקע העיקרי בכל תא לשכבה רסטורית.

רזולוציה של 100 על 100 מטרים עבור המיפוי נבחרה לאור השיקולים הבאים:

א. ביצועי מחשב: השכבה בה ניתן מזהה ייחודי לכל תא שטח היא שכבה וקטורית. מספר התאים בה הוא 2,235,946. אם לדוגמה היה המיפוי נעשה ברזולוציה של 50 על 50 מטרים מספר התאים היה מוכפל פי ארבע והשכבה היתה מאד "כבדה".

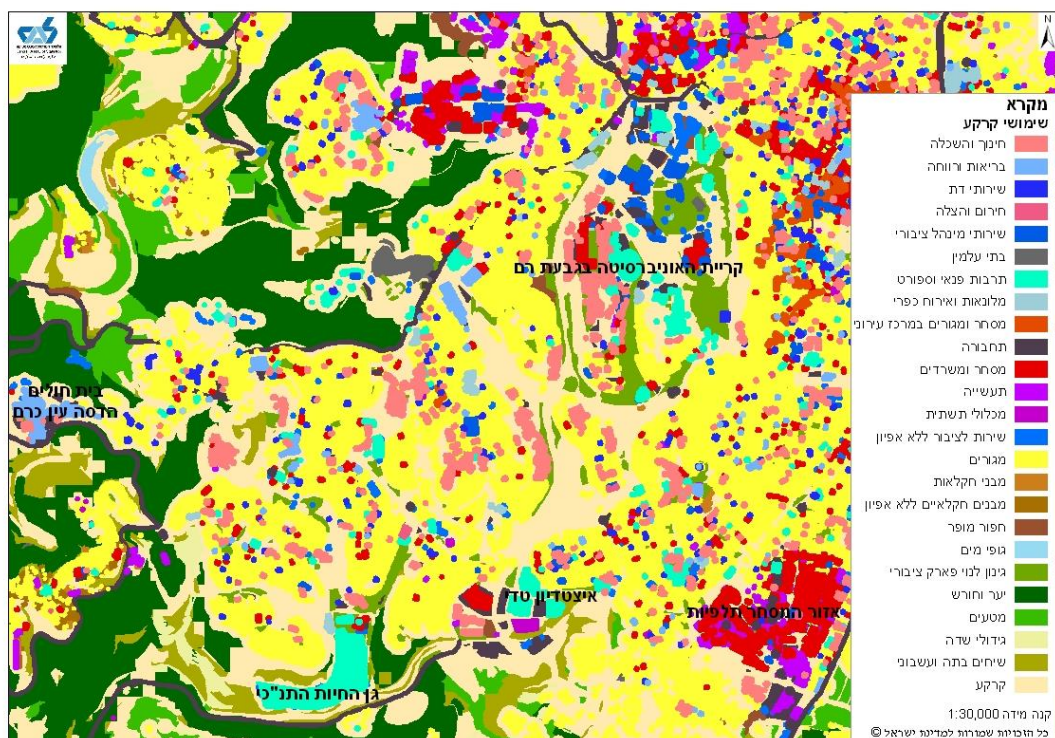
ב. מהימנות ודיוק הנתונים: קיים פער בין האיכות והדיוק של המידע שקיים עבור יישובים "קטנים" לבין מידע הקיים עבור יישובים "גדולים". בחירה ברזולוציה של 100 על 100 מטרים מאפשרת לגשר על הפער.

ג. המטרה לשמה נועדה השכבה: השכבה היא שכבה "סטטיסטית" כלומר: היא מציגה שימוש קרקע עיקרי לתא שטח. היא לא נועדה לאפשר זיהוי של כל בית ספר או כל מרכז מסחרי.

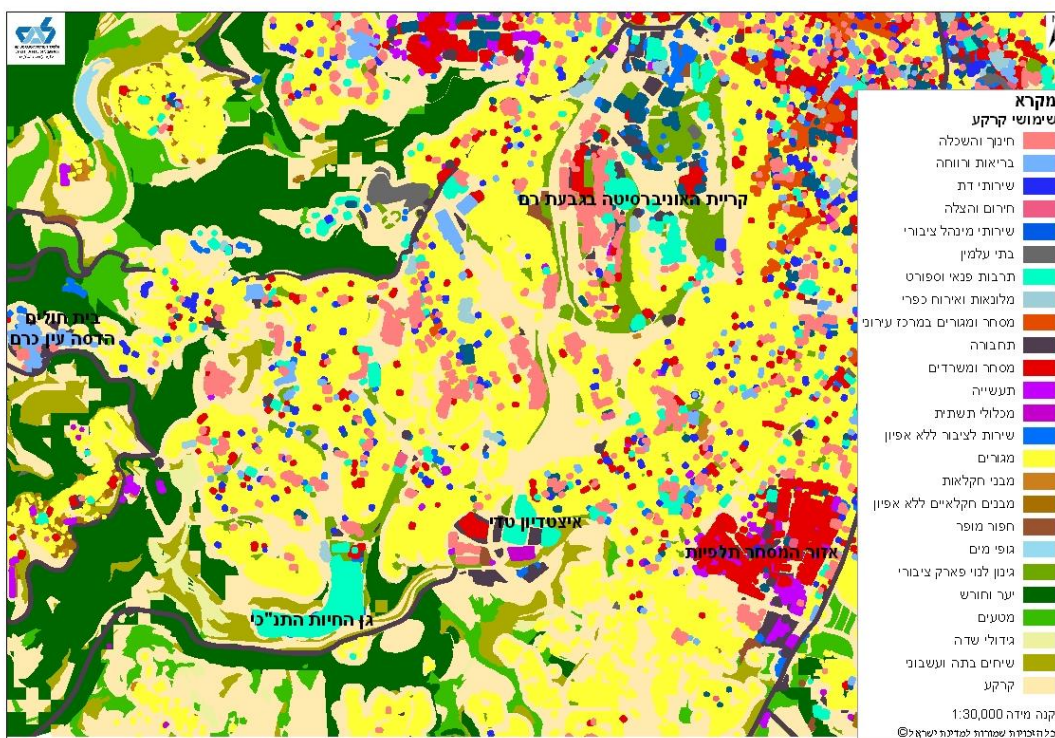
²¹ מערכות מ"ג וקטוריות מחלקות את המידע הגאוגרפי למידע נקודתי (כגון בארות או עמודי חשמל). מידע קווי (כגון נחלים, דרכים) ומידע פוליגוני (כגון תחומי שיפוט). לכל יישוב (נקודתית, קווית או פוליגונית) ישנה שורה בבסיס הנתונים של השכבה, וניתן להצמיד אליה מידע אלפא-נומרי (כגון ספיקת הבאר, שנת הסלילה של הכביש או מספר התושבים בשטח הבנוי).

²² שיטה לאחסון מידע גאוגרפי, לעיבודו ולהצגתו. האזור הנחקר מחולק לשורות ולטורים היוצרים מבנה של רשת (גריד). בכל תא ברשת שמור ערך המאפיין את תא השטח. המידע מיוצג על ידי תמונה (רסטר) בה כל פיקסל מייצג עוצמה של התופעה.

איור 2 דוגמה לשכבה הוקטורית

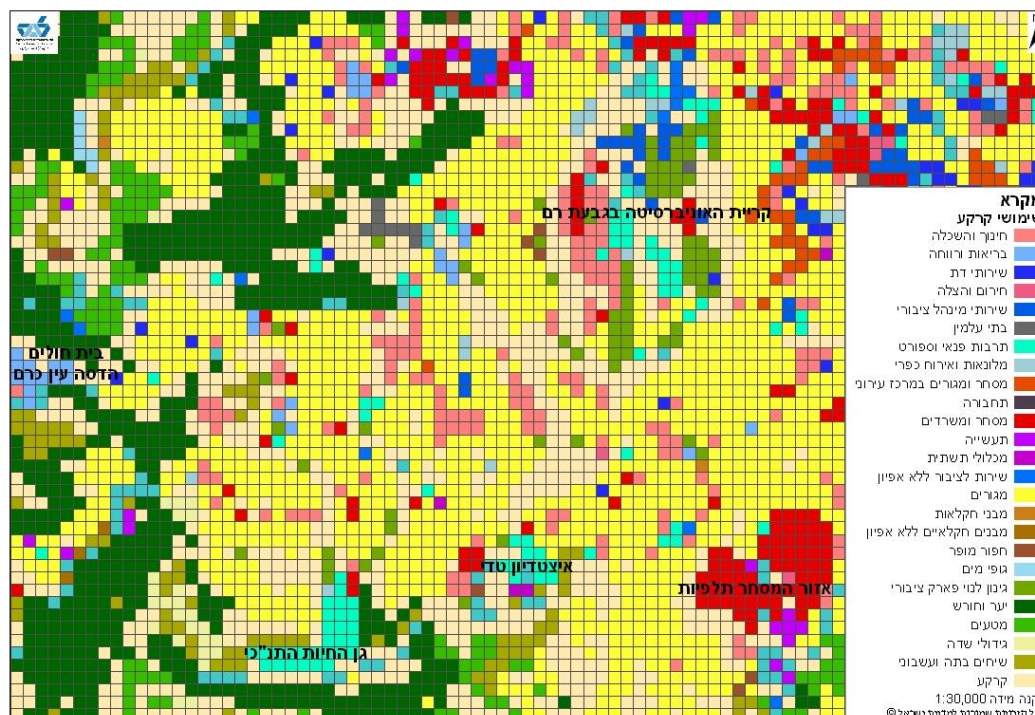


איור 3 דוגמה לשכבה הרסטריט – גודל תא 0.5 על 0.5 מטרים²³



²³ כיוון שהרזולוציה גבוהה (גודל תא 0.5 על 0.5 מטרים) איור 3 נראה זהה לאיור 2.

איור 4 דוגמה לרשת – גודל תא 100 על 100 מטרים עם ערך ה-majority של שימוש הקרקע בכל תא שטח



4.9 שכבת שימושי - הקרקע התוצר הסופי

התוצר הסופי כולל שתי שכבות מידע גאוגרפיות. האחת: שכבה רסטריאלית ארצית עם ערך שימוש הקרקע העיקרי בכל תא והשניה: שכבה וקטורית נקודתית של המרכזים של תאי השטח עם נתון על שימוש הקרקע, מספר מזהה ייחודי לכל תא וקואורדינטות של מרכז תא השטח. בנספח 1 מצוי תיאור מפורט של השכבות הללו: מבנה השדות, קובצי lyr עם ערכים לצורך הצגת השכבה לפי שלוש סימבולוגיות²⁴ שונות ופרטים נוספים.

כמו כן פורסמו קבצים עם נתונים אלפא-נומריים על שימושי הקרקע לפי רשויות מקומיות 2014²⁵, נתונים ברמה הארצית ונתונים לפי מחוזות²⁶.

בעבודה שתתפרסם בשנת 2017 (אברהם, בהכנה) יינתן פירוט רב על שיטת העבודה וכן תינתן סקירה של ממצאים בנושא שימושי הקרקע.

אנו רואים בסטטיסטיקה של שימושי קרקע תוצר סטטיסטי בעל משמעות רבה ולפיכך שוקלים אפשרות להכנת סטטיסטיקה בנושא במרווחי זמן שישקפו שינויים מהותיים בתכולה

²⁴ הכוונה היא לצבעים לפיהם ניתן להציג את השכבה לפי מספר הקבוצות של שימושי הקרקע.

²⁵ ראו קובץ "הרשויות המקומיות 2014", גיליון: "שימושי קרקע 2013", פירוט מלא" http://www.cbs.gov.il/publications16/local_authorities14_1642/excel/p_libud.xls

²⁶ ראו <http://www.cbs.gov.il/shnaton66/np1.pdf> בעמוד 9 ובעמוד 13.

המשקפים שינויים שחלו בפועל בשטח. פרק זמן זה עשוי להיות אחת לחמש שנים או כדומה.

5. סקירת ספרות

אימות נתוני מיפוי חיוני הן עבור יוצר המידע והן עבור המשתמשים בו. כדי שניתן יהיה להסתמך על הנתונים לצורך תכנון וקביעת מדיניות במגוון נושאים יש לדעת תחילה מה מידת המהימנות שלהם. מיפוי שלא עבר תהליך אימות שקול להשערה שלא נבחנה (Strahler et al., 2006).

5.1 סוגיות מרכזיות שניתן לבחון בנוגע למיפוי

1. עבור כמה אחוזים מהשטח שמופה יש התאמה בין המיפוי למציאות?
2. מהו הדפוס של הטעויות? כלומר: איזה קטגוריות לא מופו נכון ומהן הקטגוריות לפיהן הן מופו בטעות.
3. מה מידת החומרה שמייחס המשתמש במפה לטעויות שונות? במילים אחרות: כיצד מדרג המשתמש את הטעויות מבחינת חומרתן (Rossiter, 2004). בעבודה זו יינתן מענה לשתי השאלות הראשונות.

5.2 השלבים העיקריים בתהליך אימות נתוני מיפוי

5.2.1 יצירת המדגם

5.2.1.1 בחירה ביחידת הדגימה

יחידת הדגימה עשויה להיות נקודה, פיקסל או פוליגון (שטח). קיימות דוגמאות למחקרים בהם דגמו נקודות (Rozenstein & Karnieli, 2011), למחקרים בהם דגמו פיקסלים (מזור, סבוראי ובר (קותיאל), 2005) ולמחקרים בהם דגמו פוליגונים (Berghout, Robinson, & Owen, 2008; Lowry et al., 2005).

5.2.1.2 בחירה בסוג המדגם

קיימים סוגים שונים של מדגמים מתוכם ניתן לבחור במדגם שישמש לצורך בדיקת נתוני המיפוי. דוגמאות:

5.2.1.2.1 דגימה אקראית פשוטה: לכל יחידת דגימה יש סיכוי זהה להיכלל במדגם.

5.2.1.2.2 דגימת שכבות: בדוגמה של מיפוי שימושי קרקע, בדגימה אקראית פשוטה לא מובטח ייצוג לכל אחד משימושי הקרקע או לכל אחד מהאזורים הגאוגרפיים. כדי להבטיח ייצוג של שימושי קרקע קטנים או ייצוג של אזורים גאוגרפיים שונים ניתן להשתמש בדגימת שכבות. בדגימת שכבות מבטיחים מכסה מסוימת של נדגמים (לדוגמה פיקסלים) מכל שימוש קרקע או מכל אזור גאוגרפי.

ניתן להגדיר כל שימוש קרקע כשכבה ולקחת מדגם מקרי מכל שימוש קרקע. ניתן לחלק את השטח לאזורים גאוגרפיים, כל אזור גאוגרפי יוגדר כשכבה לצורך הדגימה ויילקח מדגם מקרי מכל אחד מהאזורים הגאוגרפיים.

5.2.1.2.3 דגימה שיטתית: ניתן לדוגמה לחלק את השטח למשבצות ולדגום את הנקודות המצויות במרכז כל משבצת. דהיינו: דגימה של נקודות המצויות במרחק שווה זו מזו. 5.2.1.2.4 דגימת אשכולות: השטח שמופה מחולק לאשכולות. נערכת דגימה אקראית של אשכולות ולאחר מכן נערכת דגימה אקראית של פרטים, למשל של פיקסלים, מתוך כל אשכול שנבחר. (Strahler et al., 2006).

לכל אחד מסוגי המדגמים יש יתרונות וחסרונות. בדגימה אקראית פשוטה למשל, יתקבל פחות מידע על שכבות (לדוגמה: שימושי קרקע) קטנות. יתכן שדווקא למידע על שכבה קטנה (שימוש קרקע בעל שטח קטן) יש חשיבות רבה.

דגימת שכבות מספקת מידע על כל אחת מהשכבות אולם ניתן לבצע אותה רק לאחר סיום תהליך המיפוי. כתוצאה מכך, בפרק הזמן שעובר בין המיפוי לדגימה המצב בשטח עלול להשתנות (Congalton, 1991). לדוגמה, שטח שמופה כ"מבנים חקלאיים" ובזמן שעבר בין המיפוי לעריכת המדגם הוא הפך לשטח מגורים.

יתרונה של דגימת אשכולות לדוגמה הוא הפחתת עלויות באיסוף נתונים לאימות המיפוי. כאשר נערכת בדיקה של הנתונים בשטח, דגימת אשכולות תפחית את זמן הנסיעה ממקום למקום. מצד שני, החיסרון הוא שעלולה להיות קורלציה בין הפיקסלים, המצויים באותו אשכול, מבחינת נכונות המיפוי שלהם או חוסר הנכונות של המיפוי שלהם (Strahler et al., 2006).

במחקר קיימות דוגמאות לבחירה במדגמים שונים לצורך אימות נתוני מיפוי. לדוגמה, מחקרים בהם כפי שנעשה במאמר זה, נבחרו באופן אקראי מספר קבוע של יחידות דגימה מכל אחת מהשכבות - קטגוריות (מזור ואחרים, 2005 ; Rozenstein & Karnieli, 2011). דוגמה אחרת היא מדגם שנלקח לצורך הערכה של מיפוי שימושי קרקע באוסטרליה. מספר יחידות הדגימה מכל שכבה (קטגוריה) הותאם למספר הפוליגונים, שמופו לפי שכבה זו ולשטח היחסי של השכבה מתוך סך השטח שמופה (Berghout et al., 2008).

5.2.2 בחירה במקורות מידע שישמשו לצורך אימות הנתונים

דוגמאות למקורות מידע לצורך אימות הנתונים: סקר בשטח, תצלומי אוויר ושכבות מידע גאוגרפיות.

5.2.3 הערכה של טיב המיפוי על ידי השוואה בין המדגם שנלקח מהמיפוי ובין

מקורות האימות וכימות הטעויות באמצעות מדדים סטטיסטיים

5.2.3.1 מדדים לאימות נתוני מיפוי

שני המדדים הנפוצים ביותר לצורך אימות נתוני מיפוי של כיסוי קרקע הם מדד הדיוק הכולל (Overall Accuracy) ואינדקס Kappa (Strahler et al., 2006).

הדיוק הכולל של המיפוי הוא מדד לאחוז המקרים (יחידות הדגימה) בהם יש התאמה בין המיפוי למקורות אימות.

אינדקס Kappa הוא מדד להערכת השוני שבין התאמה אמיתית של מיפוי (כנגד מפת אימות) ובין התאמה אקראית (Adam, Mutaga, Odini, & Abdel-Rahman, 2014).

מדדים אלה מחושבים לפי נתונים במטריצת עירוב (confusion matrix). זו טבלה המאפשרת להשוות בין סיווג לקטגוריות באמצעות מודל ובין נתונים ידועים (RS/GIS (Laboratories Utah State University, 2003).

הנוסחאות לפיהן מחושבים המדדים הללו והסברים מפורטים יותר עליהם, על מטריצת העירוב ועל מדדים נוספים יינתנו בפרק 6.

5.2.3.2 מבחנים להשוואה בין דיוק של שתי שיטות מיפוי

כדי להשוות בין שתי שיטות לסיווג או מיפוי של נתונים, ניתן לבחור בשיטה, עבורה הערך של מדד הדיוק הכולל היה גבוה יותר או בשיטה עבורה הערך של אינדקס Kappa היה גבוה יותר (Cingolani, Renison, Zak, & Cabido as cited in Rozenstein & Karnieli, 2011). אולם, כפי שמציינים Rozenstein & Karnieli (2011), לא כל הבדל משמעותי מבחינה סטטיסטית (מובהק סטטיסטית). לכן יש להשתמש במבחנים, הבודקים האם ההבדל בין ערכי המדדים עבור הדיוק של כל אחת משיטות המיפוי משמעותי מבחינה סטטיסטית.

דוגמאות למבחנים הבודקים האם ההבדל בין הדיוק של שתי שיטות מיפוי מובהק סטטיסטית:

5.2.3.2.1 מבחן Z (Z-test)

הנוסחה למבחן Z תינתן על ידי:

$$Z = \frac{\hat{k}_1 - \hat{k}_2}{\sqrt{\hat{\sigma}_1 + \hat{\sigma}_2}} \quad (1)$$

כאשר $\hat{k}_1$ ו- $\hat{k}_2$ מייצגים את האומדן לאינדקס Kappa של שיטת המיפוי הראשונה ושל שיטת המיפוי השנייה, בהתאמה. $\hat{\sigma}_1$ ו- $\hat{\sigma}_2$ מייצגים את האומדן לשונות של $\hat{k}_1$ ו- $\hat{k}_2$ בהתאמה (Jiang & Liu, 2011).

ההבדל בדיוק בין שתי השיטות מובהק ($p \leq 0.05$) אם הערך של Z גדול מ-1.96 (Adam et al., 2014).

5.2.3.2.2 מבחן מקנמר (McNemar's test)

אין אפשרות להשתמש במבחן Z אם אותו מדגם שימש לצורך בחינת הדיוק של שתי שיטות מיפוי (Foody as cited in Manadhar, Odeh, & Ancev, 2009).

כאשר אותו מדגם משמש לבחינת הדיוק של שתי שיטות לסיווג נתונים, ניתן להשתמש במבחן מקנמר. זהו מבחן א-פרמטרי המבוסס על חי בריבוע (chi-square), עם דרגת חופש אחת. הוא ניתן על ידי:

$$x^2 = \frac{(f_{12} - f_{21})^2}{f_{12} + f_{21}} \quad (2)$$

כאשר f_{12} מייצג את מספר המקרים שלא מופו נכון בשיטה הראשונה ומופו נכון בשיטה השנייה ו- f_{21} מייצג את מספר המקרים שמופו נכון בשיטה הראשונה ולא מופו נכון בשיטה השנייה (Manadhar et al., 2009).

חלוקת תהליך אימות הנתונים לשלושת השלבים הללו, המוצגים בסעיפים 5.2.1, 5.2.2 ו-5.2.3, מופיעה אצל Stehman and Czaplewski (1998).

5.3 שיטות "איכותיות" (qualitative) לאימות נתונים

בסעיף 2 נדונו שיטות כמותיות לאימות נתוני מיפוי. הווה אומר: שיטות שנועדו לכמת את ההתאמה בין נתוני המיפוי ובין נתוני האימות באמצעות מדדים סטטיסטיים.

קיימת אפשרות, שתתקבל התאמה "טובה" בין נתוני המיפוי ובין נתוני האימות לפי מדד סטטיסטי נתון (לדוגמה מדד הדיוק הכולל) אולם באחד מהאזורים הגאוגרפיים יהיו טעויות. כדי למנוע זאת, ניתן להשתמש הן במהלך תהליך המיפוי והן לאחר סיומו בשיטות "איכותיות" כלומר: שיטות לא כמותיות (Strahler et al., 2006).

דוגמאות לשיטות איכותיות בהן השתמשו נוסף על שיטות כמותיות הן השוואת המיפוי למפות אחרות וכן בדיקה של המיפוי על ידי "מומחים לנושא שמופה" או של "אנשים שמכירים את האזור" (Berghout et al., 2008; Lowry et al., 2005).

לצורך הערכת טיב המיפוי של שכבת שימושי קרקע 2014 בוצעו הן שיטות כמותיות באמצעות שימוש במדדים סטטיסטיים והן שיטות איכותיות באמצעות השוואה של השכבה לשכבות מידע גאוגרפיות של גופים מחוץ ללשכה המרכזית לסטטיסטיקה.

6 מתודולוגיה וממצאים

המיפוי התבסס על עיבוד מידע ממרשמים מנהליים ומשכבות מידע גאוגרפיות. בסיום תהליך העבודה נבנתה רשת ארצית רציפה של תאי שטח שגודלם 100 על 100 מטרים. לכל תא הוענק הערך של שימוש הקרקע העיקרי המצוי בו. דהיינו: שימוש הקרקע, שהשטח שלו הוא הגדול ביותר מבין שימושי הקרקע המצויים בתא.

תהליכי הבקרה נועדו לבחון באיזו מידה קיימת התאמה בין המיפוי בשכבת שימושי קרקע 2014 ובין מידע ממקורות אחרים אשר שימשו לאימות תוצאות המיפוי. יחידת השטח, לפיה

בוצעה ההשוואה, היא תא שטח בגודל 100 על 100 מטרים. זאת כיוון, שהמידע על שימושי הקרקע נתון בשכבה לפי תאי השטח הללו.

ראשית, בוצעו בקורות על שתיים עשרה קטגוריות של שימושי קרקע²⁷. הקטגוריות הן: מתחמים בנויים, מכלולי תשתית, מבנים חקלאיים, חפור מופר, גופי מים, גיבון לנוי/פארק ציבורי, יער וחורש, מטעים, גידולי שדה, שיחים בתה ועשבוני, אזור סלעי וקרקע.

קטגוריות אלה כללו את כל שימושי הקרקע וכיסויי הקרקע שמופו בשכבה. חלק מהקטגוריות מייצגות שימוש או כיסוי קרקע אחד וחלקן מורכבות ממספר שימושי קרקע.

בנספח 3 מוצג, עבור כל אחד משימושי הקרקע בשכבה מידע על הקטגוריה אליה הוא שויך לצורך הבקורות. מצוי שם גם נתון על מספר התאים, שמופו בשכבה מכל אחד משימושי הקרקע.

לאחר מכן, נערכו תהליכי בקרה על אחד עשר שימושי קרקע נבחרים המצויים בשטח הבנוי. השימושים היו: חינוך והשכלה, בריאות ורווחה, שירותי דת, חירום והצלה, שירותי מינהל ציבורי, בתי עלמין, תרבות פנאי וספורט, מלונאות ואירוח כפרי, מסחר ומגורים במרכז עירוני, מסחר ומשרדים ותעשייה.

שתיים עשרה הקטגוריות שנבדקו מייצגות בעיקר כיסויי קרקע. הן נבדקו מול פיענוח ויזואלי של אורתופוטו²⁸. הבחינה של שימושי הקרקע הבנויים התבססה גם על מקורות נוספים, שיפורטו בהמשך, שהרי לא ניתן לאמת באמצעות פיענוח ויזואלי של אורתופוטו שימושי קרקע כגון: "מסחר ומגורים במרכז עירוני" או "שירותי דת".

להלן תובא סקירה של תהליכי הבקרה שבוצעו על שתיים עשרה הקטגוריות של שימושי הקרקע. לאחר מכן תינתן סקירה של הבקורות שבוצעו על אחד עשר שימושי קרקע נבחרים המצויים בשטח הבנוי. אותם תהליכי בקרה נערכו אפוא על שתי הקבוצות הללו. ההבדל הוא במקורות המידע ששימשו לצורך אימות נתוני המיפוי.

המדדים שיוצגו בעבודה זו לצורך הבקורות הם:

1. דיוק כולל (Overall Accuracy)
2. דיוק כולל משוקלל
3. אינדקס Kappa (Cohen's Kappa Index)
4. אינדקס Kappa משוקלל (Weighted Kappa) – שקלול לפי האחוז של הקטגוריה מהסה"כ²⁹

²⁷ הקטגוריות מבוססות על רמת הפירוט הרביעית של התקן הארצי למיפוי תכסיות ושימושי קרקע. "מטעים" ו"שדות" נבדקו בנפרד לפי רמת הפירוט החמישית של התקן וגופי המים המלאכותיים והטבעיים אוחדו לצורך הבקרה.

²⁸ אורתופוטו הוא צילום אוויר סרוק ממנו סולקו העיוותים הגיאומטריים שנגרמו על ידי הטופוגרפיה והוא עוגן לרשת קואורדינטות (לוי, 2003).

הבדיקה בוצעה לפי אורתופוטו משנת 2014 ברזולוציה של 0.5 על 0.5 מטרים לפיקסל. גיליונות egd_nw ו-egd_sw היו חתוכים בחלק מהאזור בו צריך להופיע ים המלח ולכן עבור 19 תאים שהיו שם נערכה הבדיקה לפי Google Earth.

²⁹ אין הכוונה לאינדקס Kappa משוקלל לפי דירוג חומרת הטעויות אלא מדובר בשקלול לפי האחוז של הקטגוריה מהסה"כ שפותח בעבודה זו.

5. דיוק יצרן (Producer's Accuracy)

6. דיוק צרכן (User's Accuracy)

7. אינדקס Kappa עבור קטגוריה נתונה (Conditional Kappa)

6.1 בקורות על מיפוי קטגוריות שימושי הקרקע

שלב ראשון: הכנת תשתית לחישוב המדדים על קטגוריות שימושי הקרקע

בטבלה 2 מוצג עבור כל אחת משתי עשרה הקטגוריות, נתון על מספר תאי השטח (גודל תא 100 על 100 מטרים) שמופא לפיה בשכבה. הטבלה מספקת גם מידע על האחוז של התאים מקטגוריה נתונה מסך התאים בשכבת שימושי הקרקע. לצורך הבדיקה, נלקח מכל קטגוריה מדגם מקרי של 100 תאי שטח (סה"כ 1,200 תאי שטח). בסעיף 6.1.2.2 יוסבר מדוע נלקח מדגם קבוע מכל שימוש קרקע ולא מדגם פרופורציוני לגודל הקטגוריה.

כל תא שטח שנדגם, נערכה עבורו השוואה בין שימוש הקרקע לפיו הוא מופא בשכבה ובין שימוש הקרקע שניתן לו לפי פיענוח ויזואלי של אורתופוטו.

טבלה 2: התפלגות קטגוריות שימושי הקרקע, לפיהן בוצעה בקרת התוצר

קטגוריה	מספר תאי שטח בשכבה (100 על 100 מטרים)	אחוזים
מתחמים בנויים	107,861	4.82
מכלולי תשתית	18,061	0.81
מבנים חקלאיים	15,481	0.69
חפור מופר	17,079	0.76
גופי מים	57,184	2.56
גינן לנוי/פארק ציבורי	2,231	0.10
יער וחורש	158,727	7.10
מטעים	110,654	4.95
גידולי שדה	316,810	14.17
שיחים בתה ועשבוני	118,626	5.31
אזור סלעי	592,064	26.48
קרקע	721,168	32.25
סך הכל	2,235,946	100.00

המיפוי של כל תא שטח, שנדגם משכבת שימושי הקרקע, השווה לתוצאות הפיענוח הויזואלי של האורתופוטו באמצעות מטריצת עירוב (confusion matrix)³⁰. מטריצת עירוב היא טבלה, המאפשרת להשוות בין סיווג לקטגוריות באמצעות מודל ובין

³⁰ ידועה גם בשם "Error matrix" (מטריצת שגיאות) וכן בשם "מטריצת הפסד".

נתונים ידועים. השימוש במטריצת עירוב נפוץ לצורך הערכת דיוק של מודל RS/GIS (Laboratories Utah State University, 2003).

בטבלה 3 מוצגת מטריצת עירוב, שחושבה במטרה לבדוק התאמה בין המיפוי בשכבה ובין פיענוח ויזואלי של האורתופוטו (מפת האימות). המידע בשורות תואם למיפוי של שימושי הקרקע בשכבה והמידע בטורים תואם לסיווג של שימושי הקרקע לפי הפיענוח הויזואלי של האורתופוטו. הנתון בכל תא במטריצה הוא תוצאה של הצלבה בין סיווג שימושי הקרקע לפי המיפוי בשכבה ובין סיווג שימושי הקרקע לפי הפיענוח הויזואלי.

בכל תא במטריצה מוצג מספר תאי השטח התואמים להצלבה נתונה. לדוגמה, מתוך התאים שנדגמו, 95 תאים מופו בשכבה לפי הקטגוריה "מתחמים בנויים" וסווגו כ"מתחמים בנויים" גם לפי הפיענוח הויזואלי. לעומת זאת, תא אחד מהתאים שנדגמו מופה בשכבה לפי הקטגוריה "מתחמים בנויים" וסווג לפי הפיענוח הויזואלי כ"מכלולי תשתית" וכן לגבי ארבעה תאים נוספים.

טבלה 3: מטריצת עירוב של מיפוי קטגוריות שימושי הקרקע בשכבה

מספר תאי שטח בגודל 100 על 100 מטרים

פיענוח ויזואלי של אורתופוטו													
סך הכל	קרקע	אזור סלעי	שיחים בתה ועשבוני	גידולי שדה	מטעים	יער וחורש	גינן לנוי פארק ציבורי	גופי מים	חפור מופר	מבנים חקלאיים	מכלולי תשתית	מתחמים בנויים	קטגוריה
100	1	0	1	0	1	0	0	0	0	1	1	95	מתחמים בנויים
100	3	0	2	0	1	1	0	0	1	0	90	2	מכלולי תשתית
100	2	0	0	1	9	0	1	0	0	84	0	3	מבנים חקלאיים
100	11	15	3	2	4	1	0	0	57	0	6	1	חפור מופר
100	4	0	2	0	0	0	0	94	0	0	0	0	גופי מים
100	0	0	0	1	0	1	84	0	1	1	9	3	מיפוי בשכבת שימושי קרקע גינן לנוי פארק ציבורי
100	0	4	0	10	1	85	0	0	0	0	0	0	יער וחורש
100	1	0	1	1	97	0	0	0	0	0	0	0	מטעים
100	14	0	5	59	20	0	0	0	0	1	1	0	גידולי שדה
100	8	0	85	0	0	4	0	0	0	0	2	1	שיחים בתה ועשבוני
100	16	84	0	0	0	0	0	0	0	0	0	0	אזור סלעי
100	79	4	7	2	1	5	0	0	0	0	1	1	קרקע
1,200	139	107	106	76	134	97	85	94	59	87	110	106	סך הכל

שלב שני: חישוב המדדים להערכת מיפוי קטגוריות שימושי הקרקע

6.1.1 הדיוק הכולל של המיפוי (Overall Accuracy)

6.1.1.1 דיוק כולל

הדיוק הכולל של המיפוי (Overall Accuracy) הוא מדד לאחוז תאי השטח בשכבה שמופּו נכון. הקביעה מהו מספר התאים שמופּו נכון מבוססת על מספר ההסכמות/התאמות, שנמצאו במדגם בין המיפוי ובין מקורות האימות (Rossiter, 2004).

הדיוק הכולל מחושב על סמך מטריצת עירוב על ידי חלוקה של סך המקרים, המצויים באלכסון הראשי של המטריצה, בגודל המדגם (Ballestore & Qui, 2012). הדיוק הכולל חושב על ידי חלוקה של מספר תאי השטח, בהם היתה התאמה בין המיפוי לפיענוח הויזואלי, בסך תאי השטח שנדגמו. בטבלה 3 מספר תאי השטח שנדגמו הוא 1,200.

התאים במטריצה, בהם נספרו תאי שטח, שהמיפוי שלהם בשכבה תואם לפיענוח הויזואלי מצויים לאורך האלכסון הראשי של המטריצה. דהיינו, באלכסון שמתחיל מהפינה השמאלית העליונה של המטריצה ומסתיים בפינה הימנית התחתונה (החל מהמשבצת בה יש התאמה בין "מתחמים בנויים" לפי המיפוי ל"מתחמים בנויים" לפי הפיענוח הויזואלי ועד המשבצת בה יש התאמה בין "קרקע" לפי המיפוי ל"קרקע" לפי הפיענוח הויזואלי). נסמן ב- N את סך תאי השטח שנדגמו וב- Nt את מספר תאי השטח בהם היתה התאמה בין המיפוי ובין הפיענוח הויזואלי. הדיוק הכולל יסומן על ידי P . אומדן לדיוק הכולל חושב על סמך מטריצת העירוב המופיעה בטבלה 3 לפי נוסחה (3). נוסחה זו מוצגת אצל Rossiter (2004).

$$\hat{P} = \frac{Nt}{N} \quad (3)$$

נמצא כי האומדן לדיוק הכולל של מיפוי הקטגוריות, לפי נתוני המדגם המופיעים בטבלה 3 הוא 82.75%. תוצאה זו מעידה על דיוק גבוה שמשמעותו אחוז גבוה של תאי שטח שמופּו נכון.

6.1.1.2 דיוק כולל משוקלל

מכל אחת מהקטגוריות של שימושי הקרקע נלקח כאמור מדגם מקרי של 100 תאי שטח. בטבלה 2 ניתן לראות כי אחוז התאים של קטגוריה נתונה מסך כל התאים בשכבה אינו זהה בין הקטגוריות. לכן חושב דיוק כולל משוקלל, במטרה להביא בחשבון את אחוז ההצלחה במיפוי שהיה מתקבל אילו מדגם תאי השטח בכל שכבה (קטגוריה) היה פרופורציוני לסך כל התאים בשכבה.

הצלחה של מיפוי קטגוריה נתונה נמדדה לפי מידת ההתאמה, שנמצאה עבור הקטגוריה בין המיפוי ובין הפיענוח הויזואלי של האורתופוטו. הנתון על אחוז ההצלחה במיפוי קטגוריה נתונה מופיע בטבלה 6. הנתון על אחוז התאים של קטגוריה נתונה מסך כל התאים בשכבה מופיע בטבלה 2.

נסמן, pi – אחוז התאים של קטגוריה i של שימושי קרקע מסך כל התאים, qi – אחוז הצלחה במיפוי קטגוריה i של שימושי קרקע, r – מספר הקטגוריות ו- Poa ייצג את הדיוק הכולל המשוקלל.

אומדן לדיוק הכולל המשוקלל ניתן על ידי:

$$\hat{Poa} = \sum_{i=1}^r pi * \hat{qi} \quad (4)$$

טעות הדגימה של הדיוק הכולל המשוקלל תיוצג על ידי s . ni מייצג את גודל המדגם עבור קטגוריה נתונה. הערך של ni עבור כל אחת מהקטגוריות הוא 100. טעות הדגימה של הדיוק הכולל המשוקלל ניתנת על ידי:

$$s = \sqrt{\sum_{i=1}^r pi^2 * \frac{qi(1-qi)}{ni}} \quad (5)$$

ממצאים: האומדן לדיוק הכולל המשוקלל שחושב לפי נוסחה (4) הוא 80%. נתון זה יצא קטן במקצת מהדיוק הכולל של 82.75% שכן המיפוי היה מוצלח יותר, בדרך כלל, בקטגוריות היותר קטנות.

טעות הדגימה של הדיוק הכולל המשוקלל שחושבה לפי נוסחה (5) היא 1.81%. Ballestores & Qui (2012) מציינים כי הדיוק הכולל מוטה כלפי מעלה, משום שהוא לא "לוקח בחשבון" התאמה כתוצאה ממקריות. עלולה היתה להיווצר התאמה בין המיפוי בשכבה ובין הפיענוח הויזואלי של האורתופוטו גם אילו היו תאי השטח בשכבה מקבלים ערך של שימוש קרקע באופן מקרי. מסיבה זאת מחשבים את אינדקס Kappa עליו יוסבר בסעיף הבא.

6.1.2 אינדקס Kappa (Cohen's Kappa Index)

כאמור, עלולה היתה להיווצר התאמה בין תוצאות המיפוי של שימושי הקרקע ובין הסיווג של שימושי הקרקע על סמך פיענוח ויזואלי גם אם כל אחד מתאי השטח בשכבה היה מקבל ערך של שימוש קרקע באופן אקראי. אינדקס Kappa הוא מדד להערכת השוני שבין התאמה אמיתית של מיפוי (כנגד מפת אימות) ובין התאמה אקראית (Lillesand & Keifer בתוך: לוגסי, ליון, רמון ובן דור, 2004).

במילים אחרות: מדד להערכת השוני, בין מידת ההתאמה שהושגה בין תוצאות המיפוי לפענוח הויזואלי של האורתופוטו (מפת האימות) ובין התאמה אקראית. ישנן בספרות גם השגות על השימוש במדד Kappa ובפרט קיימת מחלוקת לגבי ההגדרה המתאימה של האומדן שהיה מתקבל בהתאמה מקרית (Jiang & Liu, 2011) אך השימוש בו עדיין נפוץ ומקובל בקרב החוקרים.

6.1.2.1 Kappa אינדקס

אינדקס Kappa מחושב על סמך נתונים המצויים במטריצת עירוב. נוסחה (6) לחישוב אינדקס Kappa מופיעה במדריך למשתמש, שפורסם בקשר למדד זה על ידי המעבדה לחישה מרחוק וממ"ג באוניברסיטת יוטה (RS/GIS Laboratories Utah State University, 2003).

נסמן ב- N את סך תאי השטח שנדגמו ומופיעים במטריצה, r הוא מספר השורות במטריצה, x_{ij} מסמן את הערך המופיע בשורה i ובטור j , x_{i+} הוא הסכום של שורה i ו- x_{+i} הוא הסכום של טור i .

אינדקס Kappa ניתן על ידי:

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})} \quad (6)$$

אינדקס Kappa הוא מדד הסכמה בין שני מקורות מידע. ערכי אינדקס Kappa נעים בין "1" ל"1-".

הערך "1" מייצג הסכמה/התאמה מושלמת, "0" תואם להסכמה אקראית וערכים שליליים מצביעים על הסכמה נמוכה יותר ממה שהיינו מצפים לקבל באופן מקרי (Viera & Garret, 2005). Viera and Garret (2005) מציגים טווח ערכים לצורך פרשנות של ערכי אינדקס Kappa. ערכים בטווח שבין 0.61 ל-0.99 נחשבים ערכי הסכמה/התאמה גבוהים מעבר להסכמה מקרית, ערכים שבין 0.21 ל-0.60 מייצגים ערכי הסכמה בינוניים וערכי אינדקס Kappa של 0.20 ומטה נחשבים ערכי הסכמה/התאמה נמוכים. בחישוב של אינדקס Kappa לפי נוסחה (6) על סמך נתוני ההתאמה בין המיפוי של תאי השטח בשכבה לבין סיווג ויזואלי לפי אורתופוטו המוצגים בטבלה 3, התקבל ערך אינדקס Kappa "גבוה" – 0.81.

6.1.2.2 אינדקס Kappa משוקלל³¹ (Weighted Kappa)

חישוב אינדקס Kappa משוקלל נדרש כיוון שהערך של אינדקס Kappa שהתקבל על ידי חישוב נוסחה (6) על הנתונים בטבלה 3 לא שווה לערך שהיה מתקבל ממדגם פרופורציוני (מייצג) לגודל הקטגוריה של תאי שטח מכלל התאים בארץ.

לא לקחנו מדגם פרופורציוני כי אז בחלק גדול משימושי הקרקע (צבועים בצבע ורוד³² בטבלה 4) לא היה צפוי להתקבל מדגם מספק בשביל ללמוד על יכולת הניבוי עבור אותם שימושי קרקע.

יתרונה של השיטה, לפיה לקחנו מדגם בגודל קבוע של 100 תאים עבור כל קטגוריה של שימוש קרקע, ללא התחשבות בגודל הקטגוריה, הוא שמצד אחד היא מאפשרת יכולת ניבוי גם עבור הקטגוריות הקטנות המסומנות בצבע ורוד בטבלה 4. כלומר: קטגוריות שאם המדגם היה פרופורציוני היו בהן מספר מועט ביותר של דגימות ולא היינו יכולים לאמוד בצורה טובה את טיב המיפוי בהן. מצד שני עבור קטגוריות של שימושי קרקע בעלות פרופורציה גבוהה מדגם של 100 הוא מספק בכך שניתן על ידי שקלול של אינדקס Kappa כפי שיוצג להלן להציג בדיוק מספק את ה-Kappa, שהיה מתקבל, במונחים של מדגם פרופורציוני. ניתן לראות זאת בטבלה 4.

³¹ כפי שיובהר בהמשך הסעיף אין הכוונה לאינדקס Kappa משוקלל לפי דירוג חומרת הטעויות אלא מדובר בשקלול לפי האחוז של הקטגוריה מהסה"כ שפותח בעבודה זו.

³² הכוונה לקטגוריות הבאות: "מכלולי תשתית", "מבנים חקלאיים", "חפור מופר", "גופי מים" ו"גינות לנוי פארק ציבורי".

טבלה 4: מטריצת עירוב של מיפוי קטגוריות שימושי הקרקע בשכבה – הצגת המדגם באופן פרופורציוני להתפלגות. מספר תאי שטח בגודל 100 על 100 מטרים

פיענוח ויזואלי של אורתופוטו													
קטגוריה	מתחמים בנויים	מכלולי תשתית	מבנים חקלאיים	חפור מופר	גופי מים	גינן לני פארק ציבורי	יער וחורש	מטעים	גידולי שדה	שיחים בתה ועשבוני	אזור סלעי	קרקע	סך הכל
מתחמים בנויים	54.95	0.58	0.58	0.00	0.00	0.00	0.00	0.58	0.00	0.58	0.00	0.58	57.84
מכלולי תשתית	0.19	8.75	0.00	0.10	0.00	0.00	0.10	0.10	0.00	0.19	0.00	0.29	9.72
מבנים חקלאיים	0.25	0.00	6.96	0.00	0.00	0.08	0.00	0.75	0.08	0.00	0.00	0.17	8.28
חפור מופר	0.09	0.55	0.00	5.20	0.00	0.00	0.09	0.36	0.18	0.27	1.37	1.00	9.12
גופי מים	0.00	0.00	0.00	0.00	28.88	0.00	0.00	0.00	0.00	0.61	0.00	1.23	30.72
מיפוי בשכבת שימושי קרקע גינן לני פארק ציבורי	0.04	0.11	0.01	0.01	0.00	1.01	0.01	0.00	0.01	0.00	0.00	0.00	1.20
יער וחורש	0.00	0.00	0.00	0.00	0.00	0.00	72.42	0.85	8.52	0.00	3.41	0.00	85.20
מטעים	0.00	0.00	0.00	0.00	0.00	0.00	0.00	57.62	0.59	0.59	0.00	0.59	59.40
גידולי שדה	0.00	1.70	1.70	0.00	0.00	0.00	0.00	34.01	100.32	8.50	0.00	23.81	170.04
שיחים בתה ועשבוני	0.64	1.27	0.00	0.00	0.00	0.00	2.55	0.00	0.00	54.16	0.00	5.10	63.72
אזור סלעי	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	266.92	50.84	317.76
קרקע	3.87	3.87	0.00	0.00	0.00	0.00	19.35	3.87	7.74	27.09	15.48	305.73	387.00
סך הכל	60.03	16.83	9.25	5.31	28.88	1.09	94.52	98.13	117.45	92.01	287.17	389.34	1200.00

הערך עבור תא נתון בטבלה 4 התקבל לאחר שבוצעו על הערך בתא התואם לו במיקום בטבלה 3 הפעולות הבאות: הכפלה ב-1200 (גודל המדגם), חלוקה ב-100 (גודל המדגם מכל קטגוריה) והכפלה באחוז של הקטגוריה מההתפלגות (מופיע בטבלה 2³³). לאחר מכן סוכמו הערכים בשורות ובטורים.

נסמן: N - סך תאי השטח שנדגמו ומופיעים בטבלה 3, r - מספר הקטגוריות של שימושי הקרקע, i - השורות בטבלה 3, j - העמודות בטבלה 3, x_{ij} - הערך המופיע בשורה i ובטור j , x_{ii} - ערך המופיע בהצטלבות בין טור ושורה שמספרם זהה - תאים המצויים באלכסון הראשי, ni מייצג את גודל המדגם עבור קטגוריה נתונה של שימוש קרקע (במקרה שלנו $ni = 100$) ו- W_i מייצג את המשקל בו הוכפל כל תא בטבלה 3. המשקל הוא האחוז של הקטגוריה מהסה"כ (מופיע בטבלה 2).

³³ לדוגמה: הערך של כל תא בטבלה 3 המצוי בשורה הראשונה (מתחמים בנויים) הוכפל ב-0.0482, הערך של כל תא בטבלה 3 המצוי בשורה השנייה (מכלולי תשתית) הוכפל ב-0.0081 וכן הלאה.

המשקל בו הוכפל הערך בכל תא בטבלה 3 הן לפני סכימת הערכים בשורות והן לפני סכימת הערכים בטורים הוא האחוז של הקטגוריה מהסה"כ לפי הקטגוריה המופיעה בשורות של הטבלה.

כמו כן נגדיר: $\tilde{X}_{ii} = \sum_{i=1}^r X_{ii} * W_i$ ו- $\tilde{X}_{+i} = \sum_{j=1}^r X_{ji} * W_i$, $\tilde{X}_{i+} = \sum_{j=1}^r X_{ij} * W_i$ אינדקס Kappa משוקלל (Weighted Kappa) ניתן על ידי:

$$Kw = \frac{N^2 \sum_{i=1}^r \frac{\tilde{X}_{ii}}{ni} - N^2 \sum_{i=1}^r \frac{\tilde{X}_{i+}}{ni} * \frac{\tilde{X}_{+i}}{ni}}{N^2 - N^2 \sum_{i=1}^r \frac{\tilde{X}_{i+}}{ni} * \frac{\tilde{X}_{+i}}{ni}} \quad (7)$$

בנספח 4 מפורטים שלבי היצירה של נוסחה (7) – אינדקס Kappa משוקלל (Weighted Kappa).

נמצא כי הערך של אינדקס Kappa המשוקלל שחושב על הנתונים בטבלה 3 לפי נוסחה (7) הוא 0.75.

ה"הפסד" בשיטה זו הוא העובדה, שלאומד החדש של ה-Kappa (ה-Kappa המשוקלל) יש שונות גבוהה מעט יותר מאשר אילו היינו מחשבים Kappa רגיל תחת מדגם פרופורציוני אך הפסד זה בטל בשישים ביחס לתועלת שהפקנו ממדגם לא פרופורציוני, כפי שהוסבר בתחילת הסעיף.

זאת אומרת, אומנם לא לקחנו מדגם פרופורציוני, אך נוכל לצפות שהערך האמיתי של אינדקס Kappa, שהיה מתקבל תחת מדגם פרופורציוני, לא סוטה מהערך שחושב על ידי נוסחה (7), המבוססת על מניפולציה על תוצאות המדגם השוויוני (100 מכל קטגוריה), ביותר מאחוזים בודדים.

לסיכום, נוכל להגיד באופן שמרני למדי, שהערך האמיתי של Kappa הוא מעל 0.7 או אפילו בין 0.75 ל-0.8. זהו ערך גבוה מאד המעיד על הצלחה של הניבוי דהיינו התאמה/הסכמה גבוהה בין המיפוי בשכבה ובין מפת האימות (האורתופוטו).

Rossiter (2004) מציג שיטה, שפותחה על ידי Cohen, לחישוב Kappa משוקלל (שקלול אחר מזה שמוצג בעבודה זו) תוך התחשבות בדירוג, שנותן יוצר המידע לטעויות לפי דרגת החומרה שהוא מייחס להן.

לדוגמה, ביחס לשכבת שימושי קרקע ניתן היה לקבוע למשל כי טעות לפיה מופו גידולי שדה במקום בו מצויים בפועל מטעים חמורה פחות מטעות לפיה מופה כביש במקום בו מצויים בפועל מטעים.

כדי לחשב אינדקס Kappa משוקלל לפי שיטה זו יש ליצור מטריצת עירוב (הפסד) בה יינתן משקל לכל תא בהתאם לדרגת החומרה שמייחס יוצר שכבת המידע הגאוגרפית לטעות.

התאים המצויים על האלכסון הראשי יקבלו ערך "0" כיוון שבהם יש התאמה מלאה בין המיפוי למקורות האימות. תאים המצויים במטריצה מחוץ לאלכסון הראשי יקבלו ערכים בטווח שבין "0" ל-"1". ככל שהערך, שיינתן לתא על ידי יוצר המידע יהיה קרוב יותר ל"0" המשמעות היא שיוצר המידע מייחס לטעות בין שתי הקטגוריות חומרה פחות גבוהה. הערך "0" מצביע על כך שמבחינת יוצר המידע שתי הקטגוריות נחשבות זהות לצורך הערכת הדיוק.

Rossiter מציין כי מטריצת דירוג הטעויות לא חייבת להיות סימטרית. לצורך הדוגמה, בשכבת שימושי קרקע ניתן לתת דרגת חומרה גבוהה יותר לטעות, לפיה מופו "מטעים" ובפועל מצויים במקום גידולי שדה מזו שתינתן לטעות, לפיה מופו בשכבה "גידולי שדה" ובפועל מצויים במקום מטעים.

דירוג טעויות באופן כזה הוא סובייקטיבי. כל אחד מהמשתמשים במידע על שימושי הקרקע מדרג בצורה שונה את חומרת הטעויות. זאת בהתאם למטרה עבורה הוא זקוק למידע. לדוגמה, משתמש הזקוק למידע על שימושי הקרקע בשטח הבנוי ייחס דרגת חומרה נמוכה לטעות, לפיה מופתה תכסית של "קרקע" במקום בו בפועל יש כיסוי סלעי. לעומת זאת, משתמש שזקוק למידע על שימושי הקרקע החקלאיים ייחס דרגת חומרה נמוכה לטעות, לפיה מופה שימוש הקרקע "תרבות, פנאי וספורט" במקום בו בפועל שימוש הקרקע הוא "חינוך והשכלה". לכן, בעבודה זו ניתנה לכל הטעויות דרגת חומרה זהה. במילים אחרות: הטעויות לא דורגו לפי חומרה.

יש להסתייג ולומר כי בנוגע לאימות הקטגוריות של שימושי הקרקע אמירה זו אינה מדויקת במלואה. כאשר למשל הוגדרה לשם הבדיקה קטגוריה של "גופי מים" הכוללת את גופי המים הטבעיים ואת גופי המים המלאכותיים המשמעות היא כי שתי הקטגוריות נחשבו זהות לצורך הערכת הדיוק. במילים אחרות: החלפה במיפוי בין "גופי מים טבעיים" ל"גופי מים מלאכותיים" לא הוגדרה כטעות. בנוגע לאימות שימושי הקרקע הבנויים אמירה זו נכונה ללא סייגים.

6.1.3 בדיקת הצלחת המיפוי עבור קטגוריה נתונה של שימוש קרקע ביחס לאימות

6.1.3.1 מדדים להערכת הדיוק של מיפוי קטגוריה נתונה של שימושי קרקע

להלן יוצגו שלושה מדדים אשר משמשים להערכת דיוק של קטגוריה של שימושי קרקע. דיוק היצרן ודיוק הצרכן דומים במהותם למדד הדיוק הכולל למעט העובדה שהם בודקים דיוק עבור קטגוריה נתונה.

6.1.3.1.1 דיוק לצורכי היצרן (Producer's Accuracy)

דיוק לצורכי היצרן הוא מדד המצביע על הסיכוי שתא שטח בנתוני האימות מקטגוריה נתונה מופה נכון. דהיינו: תא השטח מופה לפי הסיווג המצוי עבור הקטגוריה בנתוני האימות.

Congalton (1991) מציין כי הסיבה שמדד זה מכונה "דיוק יצרן" היא שהיצרן של המפה/שכבת המידע הגאוגרפית מעוניין לדעת, מה איכות המיפוי לפיה הוא מסוגל למפות אזור נתון.

דיוק היצרן מחושב על סמך הנתונים המוצגים במטריצת העירוב. הוא מתקבל על ידי חלוקה של מספר התאים מהקטגוריה בהם הסיווג תואם בין המיפוי לנתוני האימות בסך התאים מאותה קטגוריה לפי נתוני האימות, המצויים בטורים של מטריצת העירוב.

דיוק לצורכי היצרן מחושב לפי נוסחה (8). נוסחה זו מוצגת במאמרם של Liu, Frazier and Kumar (2007).

נסמן ב- P_{ii} את האחוז שמהווים תאי השטח בהם יש התאמה בין המיפוי לנתוני האימות בסיווג קטגוריה i מתוך סך המדגם וב- P_{+i} את אחוז התאים שסווגו לפי נתוני האימות כקטגוריה i בטור i מתוך סך המדגם. דיוק היצרן יסומן על ידי P_{ai} . האומדן לדיוק היצרן יינתן על ידי:

$$\hat{P}_{ai} = \frac{P_{ii}}{P_{+i}} \quad (8)$$

6.1.3.1.2 דיוק לצורכי המשתמש (User's Accuracy)

הדיוק לצורכי המשתמש במיפוי או במידע (להלן: דיוק הצרכן) הוא מדד המצביע על הסיכוי שתא שטח שסווג במפה לפי קטגוריה נתונה אכן מייצג את המצב בשטח.

Congalton (1991) מציין כי הסיבה שמדד זה מכונה "דיוק צרכן" היא שהצרכן מעוניין לדעת באיזו מידה הוא יכול לסמוך על המיפוי, כלומר: מהי מידת המהימנות של הנתונים.

דיוק היצרן מחושב גם הוא על סמך הנתונים המוצגים במטריצת העירוב. הוא מתקבל על ידי חלוקה של מספר התאים מהקטגוריה שהסיווג שלהם תואם בין המיפוי לנתוני האימות בסך התאים שמופו לפי הקטגוריה. דהיינו בסך התאים המצוי במטריצת העירוב בשורה של אותה קטגוריה.

הדיוק לצורכי הצרכן מחושב לפי נוסחה (9) להלן אשר מופיעה במאמר של Liu et al. (2007).

נסמן ב- P_{ii} את האחוז שמהווים תאי השטח בהם יש התאמה בין המיפוי לנתוני האימות בסיווג קטגוריה i מתוך סך המדגם וב- P_{i+} את סך התאים שסווגו במיפוי לפי קטגוריה i בשורה i . דיוק הצרכן יסווג על ידי U_{ai} .

האומדן לדיוק הצרכן יינתן על ידי:

$$\hat{U}_{ai} = \frac{P_{ii}}{P_{i+}} \quad (9)$$

דוגמה תאורטית להמחשת ההבדל בין דיוק יצרן לדיוק צרכן: אם דיוק היצרן עבור הקטגוריה "קרחונים" הוא 90% ודיוק הצרכן עבור קטגוריה זו הוא 80% אז היצרן יכול לטעון כי ב-90% מהפעמים בהן היו בשטח נתון קרחונים הם מופו כך. לעומת זאת הצרכן יכול לטעון שרק ב-80% מהפעמים בהן שטח מופה כקרחונים היו בו קרחונים. דוגמה זו מבוססת על דוגמה המופיעה אצל Congalton (1991) תוך שינוי הנתונים.

6.1.3.1.3 אינדקס Kappa עבור קטגוריה נתונה (Conditional Kappa)

אינדקס Kappa עבור קטגוריה נתונה (Conditional Kappa) מחושב לפי אותה לוגיקה של אינדקס Kappa, למעט העובדה ש-Conditional Kappa מחושב עבור קטגוריה נתונה. משתמשים בו על מנת להעריך את הדיוק של כל קטגוריה באופן נפרד (Liu et al., 2007). זהו מדד נוסף המופק על סמך הנתונים במטריצת העירוב.

במאמר של Liu et al. (2007) מוצגת הנוסחה עבור מדד זה:

נסמן ב- K_i את אינדקס kappa עבור קטגוריה i , P_{ii} הוא מספר תאי השטח בהם יש התאמה בין המיפוי לנתוני האימות בסיווג קטגוריה i , ב- P_{i+} נסמן את סך התאים שסווגו לפי קטגוריה i בשורה i וב- P_{+i} את סך התאים שסווגו לפי קטגוריה i בטור i .

הנוסחה ל-אינדקס kappa עבור קטגוריה נתונה (Conditional Kappa) תינתן על ידי:

$$K_i = \frac{P_{ii} - P_{i+} * P_{+i}}{P_{i+} - P_{i+} * P_{+i}} \quad (10)$$

בטבלה 5 מוצגות תוצאות החישוב של שלושת המדדים הללו עבור קטגוריות שימושי הקרקע.

טבלה 5: מדדים להערכת הדיוק של קטגוריה נתונה של שימוש קרקע

קטגוריה	דיוק יצרן Producer's Accuracy (אחוזים)	דיוק צרכן User's Accuracy (אחוזים)	אינדקס kappa עבור קטגוריה נתונה Conditional Kappa (User's)
בנוי	89.62	95.00	0.95
תשתית	81.82	90.00	0.89
מבנים חקלאיים	96.55	84.00	0.83
חפור מופר	96.61	57.00	0.55
גופי מים	100.00	94.00	0.93
גינון לנוי/פארק ציבורי	98.82	84.00	0.83
יער וחורש	87.63	85.00	0.84
מטעים	72.39	97.00	0.97
גידולי שדה	77.63	59.00	0.56
שיחים בתה ועשבוני	80.19	85.00	0.84
אזור סלעי	78.50	84.00	0.82
קרקע	56.83	79.00	0.76

אינדקס kappa שחושב עבור קטגוריות שימושי הקרקע הוא 0.81. עבור הקטגוריות "חפור מופר" ו-"גידולי שדה" ערך ה-kappa נמוך יותר (כ-0.6). לעומת זאת, עבור שאר הקטגוריות הוא גבוה יותר: עבור "גופי מים", "מטעים" ו-"בנוי" מעל 0.9 ועבור שאר הקטגוריות מעל 0.82.

6.1.3.2 דיוק לצורכי הצרכן – המדד המרכזי שנבחר לבדיקת הצלחת המיפוי

עבור קטגוריה נתונה

ביחס לבדיקת הצלחת המיפוי של קטגוריה נתונה, המטרה המרכזית בעבודה זו היתה לבדוק באיזו מידה (אחוזים) מיפוי של קטגוריה נתונה בשכבה תואם למצב בשטח³⁴. המדד המשרת מטרה זו הוא "דיוק לצורכי הצרכן". לפיכך, הצלחה של מיפוי קטגוריה נתונה נמדדה לפי מידת ההתאמה, שנמצאה עבור הקטגוריה בין המיפוי ובין הפיענוח היוזואלי של האורתופוטו.

³⁴ המצב בשטח נבדק כאמור לפי אורתופוטו.

טבלה 6: הצלחת המיפוי של קטגוריה של שימוש קרקע ביחס לאימות

קטגוריה	מספר תאי שטח שנדגמו (100 על 100 מטרים)	מספר התאים במדגם בהם המיפוי בשכבה תואם לפיענוח ויזואלי של אורתופוטו	אחוז התאים שמופו בהצלחה מתוך התאים שנדגמו מהקטגוריה ³⁵ (דיוק צרכן)	טעות דגימה
בנוי	100	95	95	2
תשתית	100	90	90	3
מבנים חקלאיים	100	84	84	4
חפור מופר	100	57	57	5
גופי מים	100	94	94	2
גינון לנוי/פארק ציבורי	100	84	84	4
יער וחורש	100	85	85	4
מטעים	100	97	97	2
גידולי שדה	100	59	59	5
שיחים בתה ועשבוני	100	85	85	4
אזור סלעי	100	84	84	4
קרקע	100	79	79	4

מטבלה 6 ניתן להסיק על הצלחת המיפוי של כל קטגוריה ביחס לאימות. לדוגמה, עבור הקטגוריה "מתחמים בנויים" בממוצע מופו בהצלחה 95% מהתאים בשכבה. טעות הדגימה של 95% הצלחה היא 2%. לסיכום, עבור הקטגוריות "חפור מופר" ו"גידולי שדה" האמינות של המיפוי היא כ-60% בלבד. עבור שאר הקטגוריות האמינות גבוהה, לרוב היא מעל 85%.

6.2 בקורות על שימושי הקרקע הבנויים שנבחרו

את אותן הבקורות שביצענו בסעיף 6.1 על מיפוי קטגוריות שימושי הקרקע ביצענו באותה הצורה על שימושי הקרקע הבנויים שנבחרו.

שלב ראשון: הכנת תשתית לחישוב המדדים על שימושי הקרקע הבנויים שנבחרו

בטבלה 7 מוצגים אחד עשר שימושי הקרקע הבנויים עליהם נערכו הבקורות. עבור כל אחד מהשימושים מופיע נתון על מספר תאי השטח (גודל תא 100 על 100 מטרים) שמופו לפיו בשכבה. הטבלה מספקת גם מידע על האחוז של התאים של שימוש קרקע נתון מסך התאים שמופו בשכבה לפי שימושי הקרקע הללו.

³⁵ כאשר $n_i=100$ אז אחוז ההצלחה זהה למספר התאים במדגם בהם המיפוי בשכבה תואם לפיענוח הויזואלי של האורתופוטו.

טבלה 7: התפלגות שימושי הקרקע הבנויים, עליהם בוצעו תהליכי הבקרה

שימוש הקרקע	מספר תאי שטח בשכבה (100 על 100 מטרים)	אחוזים
חינוך והשכלה	4,862	19.97
בריאות ורווחה	816	3.35
שירותי דת	350	1.44
חירום והצלה	144	0.59
שירותי מינהל ציבורי	485	1.99
בתי עלמין	875	3.59
תרבות פנאי וספורט	3,557	14.61
מלונאות ואירוח כפרי	385	1.58
מסחר ומגורים במרכז עירוני	1,139	4.68
מסחר ומשרדים	4,702	19.31
תעשייה	7,034	28.89
סך הכל	24,349	100.00

לצורך הבדיקה, נלקח מדגם מקרי של 30 תאי שטח³⁶ מכל אחד משימושי הקרקע הנ"ל (סה"כ 330 תאי שטח).

אחד המקורות באמצעותו אומתו נתוני המיפוי של שימושי הקרקע הבנויים היה Google Street View (גוגל תצוגת רחוב). זהו שירות של חברת גוגל, הניתן כחלק מ-Google Maps ו-Google Earth ומאפשר לצפות בתמונות פנורמיות ממיקומים רבים ברחובות ערים רבות בעולם. השירות מאפשר למשתמש "להסתובב", באופן וירטואלי, ברחובות הערים ולהביט לכל כיוון, כאילו הוא עצמו משוטט ברחוב.

יצירת מאגר התמונות מתבצעת על ידי כלי רכב הנוסעים ברחובות הערים כשהם מאובזרים במצלמה פנורמית המצלמת תמונות ב-360 מעלות. את התמונות מחברים יחד ומעלים לאתר, וכך מאפשרים לצופה להביט לכל כיוון בו יחפוץ (גוגל תצוגת רחוב, ללא תאריך). באמצעות Google Street View נבחנו חזיתות המבנים המצויים בתא שטח נתון. על סמך הצפייה בחזיתות המבנים, המצויים בתא שטח שנבדק, נקבע שימוש הקרקע העיקרי של תא שטח זה.

להלן מספר דוגמאות: ניתן לראות לפי צילום חזיתות המבנים כי בתא מספר 4492040 ביישוב בענה מצויים מרכז מסחרי ופנצ'ריה ולכן המסקנה היא כי שימוש הקרקע בתא זה הוא מסחר.

בתא מספר 3493743 בהוד השרון מצוי מבנה ברחוב סוקולוב 3 פינת רחוב הידידות 2 ומבנה נוסף מולו. לפי צפייה בצילום חזיתות המבנים המסקנה היא כי מדובר בתא בו שימוש הקרקע הוא "מסחר ומגורים". בתא מספר 2299651 בבאר שבע נראה בצילומים המבנה

³⁶ השאיפה היא לדגימה מינימלית (חיסכון בזמן ובכסף) שעדיין מאפשרת הסקת מסקנות ברמת מהימנות סבירה.

של עיריית באר שבע ולכן נקבע כי שימוש הקרקע המייצג תא זה הוא "שירותי מינהל ציבורי". בתא 4604085 ניתן לראות, לפי צילומי הרחובות, כי מצוי שם "בית חולים לגליל המערבי נהריה" ולכן המסקנה היא כי השימוש הוא "בריאות ורווחה".

היו מקרים בהם הבדיקה ב-Google Street View העלתה כי אומנם מצויים בתא השטח מבנים, שהשימוש שלהם תואם לשימוש הקרקע, לפיו מופה תא השטח בשכבה אך אין זה שימוש הקרקע העיקרי של תא השטח. לדוגמה תא 3516179 בכפר סבא מופה לפי שימוש הקרקע "בריאות ורווחה". הצפייה בחזיתות המבנים הראתה כי אומנם מצוי בתא זה מבנה שכתובתו "רחוב החי"ל 15" והוא משמש כקופת חולים כללית אולם יתר המבנים המצויים בתא זה משמשים למגורים: מבנה אחד שהכתובת שלו "רחוב החי"ל 18" ומבנה נוסף שהכתובות שלו הן "החי"ל 20" ו"החי"ל 22". התא כולל גם חלק ממבנה שכתובתו "המעפילים 6" ומשמש אף הוא למגורים. לפי הבדיקה, השימוש העיקרי של תא שטח זה הוא "מגורים". לכן לפי הבדיקה השיוך של תא זה נקבע כ"אחר".

בתא השטח, המתואר בדוגמה זו, מבני המגורים נמצאים בחלק הדרומי של תא השטח וקופת החולים מצויה בחלק הצפוני של תא השטח. אם תאי השטח היו קטנים יותר, יתכן שחלק מהשטח היה ממופה כ"בריאות ורווחה" וחלק אחר היה ממופה כ"מגורים".

יש שימושים שקשה לזהותם באמצעות צפייה בחזיתות המבנים כמו לדוגמה בתי כנסת. הסיבה היא שלא תמיד מצוי עליהם שלט בולט הניתן לקריאה. קיימים בתי כנסת שנראים מבחוץ כמו מבני מגורים וישנם בתי כנסת המצויים במבני מגורים.

בטבלה 8 ניתן לראות כי 12 תאי שטח שמופו כ"שירותי דת" סווגו בבדיקה כ"אחר". בחלק מהמקרים השימוש העיקרי של התא לא היה "שירותי דת" ובחלק מהמקרים לא ניתן היה לזהות את בתי הכנסת לפי צילומי חזיתות המבנים.

בטבלה 7 ניתן לראות שהתאים שמופו בשכבה לפי שימוש הקרקע "שירותי דת" מהווים 1.44% בלבד מתוך תאי השטח הבנויים, שמופו בשכבה לפי אחד עשר שימושי הקרקע הבנויים שנבחרו לצורך הבקורות.

מקור מידע נוסף בו נעשה שימוש לצורך האימות הוא האורתופוטו משנת 2014. לדוגמה, כאשר מופיע באורתופוטו אצטדיון, ניתן להסיק ששימוש הקרקע הוא "תרבות פנאי וספורט". בתי עלמין ניתנים גם הם לזיהוי לפי אורתופוטו. קיימים מלונות הניתנים לזיהוי לפי אורתופוטו, בהתאם לצורת המבנה ולבריכת השחייה. לדוגמה: בתא מספר 47801 מצוי מלון ספורט קלאב באילת.³⁷

היו מקרים בהם לא ניתן היה לקבוע מהו שימוש הקרקע לפי Google Street View. למשל רחוב שמצוי בסמטה צדדית שלא צולמה או יישוב קטן שלא צולם או מקום שצולם אך לפי

³⁷ המידע כי שם המלון הוא "ספורט קלאב" מבוסס על מקורות מידע אחרים ולא על האורתופוטו.

הצילום מבחוץ לא ניתן לקבוע מהו שימוש הקרקע של המבנים. במקרים אלו גם בחינה ויזואלית של האורתופוטו לא סייעה לקבוע מהו שימוש הקרקע של תא השטח. לכן, במקרים הללו נקבע שימוש הקרקע בדרכים אחרות. לדוגמה, בירור טלפוני מול מזכירות היישוב. דוגמה אחרת, יכולה לשמש הבדיקה של תא 4871413 בקריית שמונה. בתא זה מופיע באורתופוטו מתחם עם מבנה גדול. המבנה לא צולם ב-Google Street View. תא זה מופה בשכבה לפי שימוש הקרקע "חירום והצלה" כיוון שלפי מקורות המידע, ששימשו לצורך המיפוי, מצויה במבנה זה תחנת המשטרה של קריית שמונה. לפי מיפוי הכתובות של הלשכה המרכזית לסטטיסטיקה כתובתו של המבנה היא "רחוב המכבים 8 קריית שמונה". לפי תצוגה מרחבית של שכבת המבנים ושל שכבת הרחובות, ניתן לראות כי מדרום למבנה מצוי רחוב סלינג'ר ומצפון למבנה מצוי רחוב המכבים. לצורך הבדיקה נערך חיפוש של משטרת קריית שמונה באתר של משטרת ישראל. באתר של משטרת ישראל נתונה הכתובת "המכבים 8" עבור תחנת המשטרה של קריית שמונה וכן מוצגת שם מפה עם מיקום התחנה. ניתן לראות במפה כי מדובר במבנה המצוי בתא 4871413.

מצד אחד, העובדה שלצורך האימות נערך חיפוש של תחנת משטרה (או במקרים אחרים חיפוש של רשת שיווק, מוסד חינוכי וכו') כי היה ידוע ממקורות המידע, לפיהם נערך המיפוי, שיש במקום תחנת משטרה פוגעת בחוסר התלות בין המיפוי לבין האימות אולם מצד שני, האימות בוצע על סמך מקור מידע שלא נעשה בו שימוש לצורך המיפוי (במקרה זה מידע מהאתר של משטרת ישראל).

במאמר מוסגר: כיוון שהמיפוי התבסס על מגוון מקורות מידע, היו מגוון מקורות מידע שלא ניתן היה להסתמך עליהם לצורך האימות. לא ניתן היה להסתמך לדוגמה על מידע המצוי ב-GovMap אתר המפות הרשמי של מדינת ישראל. זאת כיוון שמצויות שם שכבות הן של המרכז למיפוי ישראל והן של הלשכה המרכזית לסטטיסטיקה ומקורות המידע כללו שכבות מידע של שני הגופים. כמו כן, לא ניתן היה להסתמך על מפות המופיעות באתר Google Maps³⁸ כי מפות הרקע המופיעות שם יוצרו על ידי חברת "מפה" ומקורות המידע כללו שכבות של חברת מפה.

מתוך 330 תאי השטח של שימושי הקרקע הבנויים עליהם נעשו הבקורות 99 תאים (30%) נבדקו באמצעות Google Street View, 153 תאים (46%) נבדקו באמצעות אורתופוטו ו-78 תאים (24%) נבדקו בדרכים אחרות, כפי שפורט לעיל.

³⁸ שימוש באתר Google Maps נעשה רק לצורך העלאת תצלומי הרחוב. הללו אינם תלויי מיפוי.

טבלה 8: בדיקת התאמה בין תוצאות המיפוי לבין צילומי רחוב של Google Street

View או פיענוח ויזואלי של אורתופוטו או מקורות נוספים³⁹

מספר תאי שטח בגודל 100 על 100 מטרים

בדיקת שימוש הקרקע לפי צילומי רחוב של Google Street View ו/או לפי פיענוח ויזואלי של אורתופוטו

מיפוי בשכבת
שימושי קרקע

סך הכל	אחר	תעשייה	מסחר ומשרדים	מסחר ומגורים במרכז עירוני	מלונאות ואירוח כפרי	תרבות פנאי וספורט	בתי עלמין	שירותי מינהל ציבורי	חירום והצלה	שירותי דת	בריאות ורווחה	חינוך והשכלה	שימוש הקרקע
30	8	0	1	0	0	1	0	0	0	0	0	20	חינוך והשכלה
30	9	2	0	0	0	0	0	0	0	0	18	1	בריאות ורווחה
30	12	0	1	0	0	0	0	1	0	15	0	1	שירותי דת
30	7	0	1	0	0	1	0	3	16	0	1	1	חירום והצלה
30	5	0	3	0	0	1	0	18	2	0	0	1	שירותי מינהל ציבורי
30	1	0	0	0	0	0	29	0	0	0	0	0	בתי עלמין
30	4	1	0	0	0	24	1	0	0	0	0	0	תרבות פנאי וספורט
30	3	1	0	0	25	0	0	0	0	0	0	1	מלונאות ואירוח כפרי
30	3	0	0	26	0	0	0	0	0	0	0	1	מסחר ומגורים במרכז עירוני
30	4	0	22	0	0	1	0	1	0	0	0	2	מסחר ומשרדים
30	7	23	0	0	0	0	0	0	0	0	0	0	תעשייה
330	63	27	28	26	25	28	30	23	18	15	19	28	סך הכל

שלב שני: חישוב המדדים להערכת מיפוי שימושי הקרקע הבנויים שנבחרו

6.2.1 הדיוק הכולל של המיפוי (overall accuracy)

6.2.1.1 דיוק כולל

אומדן לדיוק הכולל של מיפוי שימושי הקרקע הבנויים, שנבחרו לצורך הבקרה הנו 71.52%. האומדן לדיוק הכולל חושב לפי נוסחה (3) על סמך הנתונים המוצגים במטריצת העירוב המוצגת בטבלה 8. הערך של N בטבלה 8 הוא 330. Nt מייצג את מספר תאי השטח בהם היתה התאמה בין המיפוי ובין מקורות האימות.

³⁹ פירוט לעיל בסעיף 6.2

6.2.1.2 דיוק כולל משוקלל

מכל אחד משימושי הקרקע הבנויים, שנבחרו לצורך הבקורות, נלקח כאמור מדגם מקרי של 30 תאי שטח. במטרה להביא בחשבון גם את אחוז ההצלחה במיפוי כל אחד משימושי הקרקע הללו וגם את אחוז התאים של כל שימוש קרקע מסך התאים, חושב דיוק כולל משוקלל.

הצלחה של מיפוי שימוש קרקע נתון נמדדה לפי מידת ההתאמה, שנמצאה עבור שימוש הקרקע בין המיפוי ובין המקורות ששימשו לאימות נתוני המיפוי.

אחוז ההצלחה במיפוי שימוש קרקע נתון מופיע בטבלה 11. אחוז התאים של כל שימוש קרקע מסך התאים בשכבה, שמופוי לפי שימושי הקרקע הללו, מופיע בטבלה 7. הדיוק הכולל המשוקלל אשר חושב לפי נוסחה (4) הוא 74%. ניתן לראות כי הדיוק הכולל המשוקלל יוצא כעת גדול מהדיוק הלא משוקלל שכן הצלחה טובה במיפוי היתה דווקא בקטגוריות הגדולות של שימושי הקרקע.

טעות הדגימה של הדיוק הכולל המשוקלל חושבה לפי נוסחה (5). גודל המדגם עבור שימוש קרקע נתון מסומן על ידי ni . הערך של ni עבור כל אחד משימושי הקרקע הללו הוא 30. טעות הדגימה של הדיוק הכולל המשוקלל היא 3.43%.

6.2.2 אינדקס Kappa (Cohen's Kappa Index)

6.2.2.1 אינדקס Kappa

אינדקס Kappa חושב לפי נוסחה (6) על סמך מטריצת העירוב המוצגת בטבלה 8. בחישוב של אינדקס Kappa על סמך נתוני ההתאמה בין המיפוי של תאי השטח בשכבה לבין סיווג תאי השטח לפי מקורות האימות (צילומי רחוב של Google Street View, פיענוח ויזואלי של אורתופוטו ומקורות נוספים⁴⁰).

ערך אינדקס Kappa שהתקבל הוא "0.69". לפי טווח הערכים⁴¹ שמציגים Viera and Garret (2005) לצורך פרשנות של ערכי אינדקס Kappa, ערך זה נחשב גבוה.

6.2.2.2 אינדקס Kappa משוקלל (Weighted Kappa)

בדומה לאמור בסעיף 6.1.2.2, הערך של אינדקס Kappa שהתקבל על ידי חישוב נוסחה (6) על הנתונים בטבלה 8 לא שווה לערך שהיה מתקבל ממדגם פרופורציוני (מייצג) של שימושי קרקע בנויים. לא לקחנו מדגם פרופורציוני כי אז בחלק גדול משימושי הקרקע

⁴⁰ בירורים טלפוניים וחיפוש באינטרנט, פירוט לעיל בסעיף 6.2.
⁴¹ פירוט של טווח הערכים מופיע בסעיף 6.1.2.1

(צבועים בצבע ורוד⁴² בטבלה 9) לא היה מדגם מספק בשביל ללמוד על יכולת הניבוי עבור אותם שימושי קרקע בנויים.

יתרונה של השיטה, לפיה לקחנו מדגם של 30 תאים עבור כל שימוש קרקע הוסבר בסעיף 6.1.2.2, והוא שהיא מאפשרת יכולת ניבוי גם עבור שימושי הקרקע המסומנים בצבע ורוד בטבלה 9. נוסף על כך, הצלחנו על ידי תיקון הנוסחה של אינדקס Kappa (שקלול נכון) להציג את ה-Kappa, שהיה מתקבל, במונחים של מדגם פרופורציוני בלי לאבד הרבה מיעילות האומדן. דהיינו: בלי להגדיל הרבה את סטיית התקן של האומדן ל-Kappa. ניתן לראות זאת בטבלה 9.

הערך עבור תא נתון בטבלה 9 התקבל לאחר שבוצעו על הערך בתא התואם לו במיקום בטבלה 8 הפעולות הבאות: הכפלה ב-330 (גודל המדגם), חלוקה ב-30 (גודל המדגם מכל שימוש קרקע) והכפלה באחוז של שימוש הקרקע מההתפלגות (מופיע בטבלה 7⁴³). לאחר מכן סוכמו הערכים בשורות ובטורים.

⁴² הכוונה לשימושי הקרקע הבאים: "בריאות ורווחה", "שירותי דת", "חירום והצלה", "שירותי מינהל ציבורי", "בתי עלמין", "מלונאות ואירוח כפרי" ו"מסחר ומגורים במרכז עירוני".

⁴³ לדוגמה, הערך בכל תא המופיע בטבלה 8 בשורה הראשונה (חינוך והשכלה) הוכפל ב-0.1997, הערך בכל תא המופיע בטבלה 8 בשורה השנייה (בריאות ורווחה) הוכפל ב-0.0335 וכן הלאה.

טבלה 9: בדיקת התאמה בין תוצאות המיפוי לבין צילומי רחוב של Google Street

View או פיענוח ויזואלי של אורתופוטו או מקורות נוספים – הצגת המדגם באופן

פרופורציוני להתפלגות

מספר תאי שטח בגודל 100 על 100 מטרים

בדיקת שימוש הקרקע לפי צילומי רחוב של Google Street View ו/או לפי פיענוח ויזואלי של אורתופוטו

סך הכל	אחר	תעשייה	מסחר ומשרדים	מסחר ומגורים במרכז עירוני	מלונאות ואירוח כפרי	תרבות פנאי וספורט	בתי עלמין	שירותי מינהל ציבורי	חירום והצלה	שירותי דת	בריאות ורווחה	חינוך והשכלה	שימוש הקרקע
65.90	17.57	0.00	2.20	0.00	0.00	2.20	0.00	0.00	0.00	0.00	0.00	43.93	חינוך והשכלה
11.06	3.32	0.74	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.63	0.37	בריאות ורווחה
4.75	1.90	0.00	0.16	0.00	0.00	0.00	0.00	0.16	0.00	2.38	0.00	0.16	שירותי דת
1.95	0.45	0.00	0.06	0.00	0.00	0.06	0.00	0.19	1.04	0.00	0.06	0.06	חירום והצלה
6.57	1.09	0.00	0.66	0.00	0.00	0.22	0.00	3.94	0.44	0.00	0.00	0.22	שירותי מינהל ציבורי
11.85	0.39	0.00	0.00	0.00	0.00	11.45	0.00	0.00	0.00	0.00	0.00	0.00	בתי עלמין
48.21	6.43	1.61	0.00	0.00	0.00	38.57	1.61	0.00	0.00	0.00	0.00	0.00	תרבות פנאי וספורט
5.21	0.52	0.17	0.00	0.00	4.35	0.00	0.00	0.00	0.00	0.00	0.00	0.17	מלונאות ואירוח כפרי
15.44	1.54	0.00	0.00	13.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.51	מסחר ומגורים במרכז עירוני
63.72	8.50	0.00	46.73	0.00	0.00	2.12	0.00	2.12	0.00	0.00	0.00	4.25	מסחר ומשרדים
95.34	22.25	73.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	תעשייה
330.00	63.97	75.61	49.81	13.38	4.35	43.18	13.06	6.42	1.48	2.38	6.70	49.68	סך הכל

מיפוי בשכבת שימושי קרקע

לסיכום, נוכל להגיד באופן שמרני למדי, שהערך האמיתי של Kappa הוא כ-0.7. זהו ערך גבוה המעיד על הצלחה של הניבוי. דהיינו: התאמה/הסכמה גבוהה בין המיפוי בשכבה ובין מקורות האימות.

6.2.3 בדיקת הצלחת המיפוי של שימוש קרקע נתון ביחס לאימות

המדדים "דיוק יצרן", "דיוק צרכן" ו-"אינדקס Kappa עבור שימוש קרקע נתון" חושבו לפי נוסחאות (8), (9) ו-(10) בהתאמה. בטבלה 10 מוצגות תוצאות החישוב של שלושת המדדים הללו עבור שימושי הקרקע הבנויים.

6.2.3.1 מדדים להערכת הדיוק של מיפוי שימוש קרקע נתון

טבלה 10: מדדים להערכת הדיוק של שימוש קרקע נתון

שימוש הקרקע	דיוק יצרן Producer's Accuracy (אחוזים)	דיוק צרכן User's Accuracy (אחוזים)	אינדקס Kappa עבור שימוש קרקע נתון Conditional Kappa (User's)
חינוך והשכלה	71.43	66.67	0.64
בריאות ורווחה	94.74	60.00	0.58
שירותי דת	100.00	50.00	0.48
חירום והצלה	88.89	53.33	0.51
שירותי מינהל ציבורי	78.26	60.00	0.57
בתי עלמין	96.67	96.67	0.96
תרבות פנאי וספורט	85.71	80.00	0.78
מלונאות ואירוח כפרי	100.00	83.33	0.82
מסחר ומגורים במרכז עירוני	100.00	86.67	0.86
מסחר ומשרדים	78.57	73.33	0.71
תעשייה	85.19	76.67	0.75

הערך של אינדקס Kappa שחושב עבור שימושי הקרקע הבנויים הוא כ-0.7. בהשוואה לערך זה, אינדקס Kappa עבור קטגוריה נתונה, שחושב עבור הקטגוריות "חינוך והשכלה", "בריאות ורווחה", "שירותי דת", "חירום והצלה" ו-"שירותי מינהל ציבורי" נמוך יותר (כ-0.6). אינדקס Kappa עבור קטגוריה נתונה, שחושב עבור הקטגוריות "בתי עלמין", "תרבות פנאי וספורט", "מלונאות ואירוח כפרי", "מסחר ומגורים במרכז עירוני", "מסחר ומשרדים" ו-"תעשייה" גבוה יותר (בין 0.71 ל-0.96) בהשוואה לערך זה.

עבור כל שימושי הקרקע, למעט "בתי עלמין" דיוק היצרן גבוה יותר מדיוק הצרכן. יחד עם זאת, כיוון שנלקח מדגם שכבות מקרי, יתכן שדיוק היצרן מוטה כלפי מעלה. זאת כיוון שאם נלקח מדגם משימוש קרקע שמופה נכון גדל הסיכוי שיהיו יותר תאים בעלי שיוך "נכון".

6.2.3.2 המדד המרכזי שנבחר לבדיקת הצלחת המיפוי עבור שימוש קרקע נתון

בדומה לאמור בסעיף 6.1.3.2, המטרה המרכזית בעבודה זו ביחס לבדיקת הצלחת המיפוי של שימוש קרקע נתון היתה לבדוק באיזו מידה (אחוזים) מיפוי של שימוש קרקע נתון בשכבה תואם למצב בשטח. המדד המשרת מטרה זו הוא "דיוק לצורכי הצרכן". לכן, הצלחה של מיפוי שימוש קרקע נתון נמדדה לפי מידת ההתאמה, שנמצאה עבור שימוש הקרקע בין המיפוי ובין המקורות ששימשו לאימות נתוני המיפוי⁴⁴.

⁴⁴ צילומי רחוב של Google Street View, פיענוח ויזואלי של אורתופוטו ומקורות נוספים, ראו הערה מספר 40.

על סמך טבלה 11 להלן ניתן להסיק לדוגמה כי עבור שימוש הקרקע "חינוך והשכלה" בממוצע מופו בהצלחה 67% מהתאים בשכבה. טעות הדגימה של 67% הצלחה היא 9%. עבור שימוש הקרקע "שירותי דת" בממוצע מופו נכון 50% מהתאים בשכבה. טעות הדגימה של 50% הצלחה היא 9%.

תאי השטח שמופו בשכבה לפי שימוש הקרקע "שירותי דת" מהווים 0.02 אחוזים בלבד מסך התאים בשכבה וזו דוגמה טובה לכך שאם המדגם היה פרופורציוני לא היינו יודעים להעריך את טיב המיפוי של שימוש קרקע זה.

טבלה 11: הצלחת המיפוי של שימוש קרקע ביחס לאימות

שימוש הקרקע	מספר תאי שטח שנדגמו (100 על 100 מטרים)	מספר התאים במדגם בהם המיפוי בשכבה תואם לסיווג התא לפי המקורות ששימשו לאימות	אחוז התאים שמופו בהצלחה מתוך התאים שנדגמו משימוש הקרקע (דיוק צרכן)	טעות דגימה
חינוך והשכלה	30	20	67	9
בריאות ורווחה	30	18	60	9
שירותי דת	30	15	50	9
חירום והצלה	30	16	53	9
שירותי מינהל ציבורי	30	18	60	9
בתי עלמין	30	29	97	3
תרבות פנאי וספורט	30	24	80	7
מלונאות ואירוח כפרי	30	25	83	7
מסחר ומגורים במרכז עירוני	30	26	87	6
מסחר ומשרדים	30	22	73	8
תעשייה	30	23	77	8

תאי השטח שמופו בשכבה לפי שימוש הקרקע "חירום והצלה" מהווים 0.01 אחוזים בלבד מסך התאים בשכבה ולכן האמור לעיל נכון גם לגבי שימוש קרקע זה. לסיכום, עבור שימושי הקרקע "שירותי דת", "חירום והצלה", "בריאות ורווחה" ו"שירותי מינהל ציבורי" אמינות המיפוי נמוכה – 60% ומטה. עם זאת, ניתן לראות בטבלה 7 כי האחוז של כל אחד משימושי הקרקע הללו מתוך תאי השטח הבנויים שנבחרו נמוך ומנספח 3 ניתן להסיק כי האחוז שמהווה כל אחד משימושי הקרקע הללו מתוך התאים בשכבה קטן אף יותר. עבור שימושי הקרקע "חינוך והשכלה", "מסחר ומשרדים" ו"תעשייה" אמינות המיפוי בינונית ועבור שימושי הקרקע "תרבות פנאי וספורט", "מלונאות ואירוח כפרי", "מסחר ומגורים במרכז עירוני" ו"בתי עלמין" אמינות המיפוי גבוהה – מעל 80%.

7. סיכום ומסקנות

7.1 סיכום של המתודולוגיה והממצאים

בסיום תהליך העבודה נבנתה רשת ארצית רציפה של תאי שטח שגודלם 100 על 100 מטרים. לכל תא הוענק הערך של שימוש הקרקע העיקרי המצוי בו. דהיינו: שימוש הקרקע, שהשטח שלו הוא הגדול ביותר מבין שימושי הקרקע המצויים בתא.

תהליכי הבקרה נועדו לבחון באיזו מידה קיימת התאמה בין המיפוי בשכבת שימושי קרקע ובין מידע ממקורות אחרים אשר שימשו לאימות תוצאות המיפוי.

יחידת השטח, לפיה בוצעה ההשוואה, היא תא שטח בגודל 100 על 100 מטרים משום שהמידע על שימושי הקרקע נתון בשכבה לפי תאי השטח הללו.

בוצעו בקורות על שתיים עשרה קטגוריות של שימושי קרקע אשר כללו את כל שימושי הקרקע וכיסויי הקרקע שמופו בשכבה. כמו כן, נערכו תהליכי בקרה על אחד עשר שימושי קרקע נבחרים המצויים בשטח הבנוי.

לצורך הבדיקה של שתיים עשרה הקטגוריות, לפיהן קובצו שימושי הקרקע, נלקח מכל קטגוריה מדגם מקרי של 100 תאי שטח (סה"כ 1,200 תאי שטח). לא נלקח מדגם פרופורציוני לגודל הקטגוריה כי אז בחלק גדול משימושי הקרקע לא היה מדגם מספק בשביל ללמוד על יכולת הניבוי עבור אותם שימושי קרקע.

כל תא שטח שנדגם, נערכה עבורו השוואה בין שימוש הקרקע, לפיו הוא מופה בשכבה ובין שימוש הקרקע, שניתן לו לפי פיענוח ויזואלי של אורתופוטו. המיפוי הושווה לתוצאות הפיענוח הויזואלי של האורתופוטו באמצעות מטריצת עירוב. נמצא כי האומדן לדיוק הכולל של מיפוי הקטגוריות הללו הוא 82.75%.

כדי להביא בחשבון גם את אחוז ההצלחה במיפוי שהיה מתקבל אילו מדגם תאי השטח בכל שכבה (קטגוריה) היה פרופורציוני לסך כל התאים בשכבה, חושב דיוק כולל משוקלל. האומדן לדיוק הכולל המשוקלל הוא 80% וטעות הדגימה שלו היא 1.81%.

חושב אינדקס Kappa - מדד להערכת השוני שבין התאמה אמיתית של מיפוי (כנגד מפת אימות) ובין התאמה אקראית. התקבל ערך אינדקס Kappa "גבוה" – 0.81.

הערך של אינדקס Kappa שהתקבל אינו שווה לערך שהיה מתקבל ממדגם פרופורציוני. לא לקחנו מדגם פרופורציוני כי אז בחלק גדול משימושי הקרקע לא היה מדגם מספק בשביל ללמוד על יכולת הניבוי עבור אותם שימושי קרקע.

יתרונה של השיטה, לפיה לקחנו מדגם בגודל קבוע של 100 תאים עבור כל קטגוריה של שימוש קרקע, ללא התחשבות בגודל הקטגוריה, הוא שהיא מאפשרת יכולת ניבוי גם עבור הקטגוריות הקטנות.

נוסף על כך, ניתן על ידי השקלול של אינדקס Kappa (לפי האחוז של הקטגוריה מהסה"כ) להציג את ה-Kappa, שהיה מתקבל, במונחים של מדגם פרופורציוני. הערך שהתקבל עבור אינדקס Kappa המשוקלל הוא 0.75.

ה"הפסד" בשיטה זו הוא העובדה שלאומד החדש של ה-Kappa (ה-Kappa המשוקלל) יש שונות גבוהה מעט יותר מאשר אילו היה מחושב תחת מדגם פרופורציוני אך הפסד זה בטל בשישים ביחס לתועלת, שהפקנו ממדגם לא פרופורציוני.

ניתן לומר באופן שמרני למדי, שהערך האמיתי של Kappa הוא מעל 0.7 או אפילו בין 0.75 ל-0.8. זהו ערך גבוה מאד המעיד על הצלחה של הניבוי דהיינו התאמה/הסכמה גבוהה בין המיפוי בשכבה ובין מפת האימות (האורתופוטו).

נבדקה גם הצלחת המיפוי של כל קטגוריה של שימוש קרקע ביחס לאימות. נמצא כי עבור הקטגוריות "חפור מופר" ו"גידולי שדה" האמינות של המיפוי היא כ-60% בלבד. עבור שאר הקטגוריות האמינות גבוהה, לרוב היא מעל 85%.

לצורך הבקורות על אחד עשר שימושי קרקע נבחרים, המצויים בשטח הבנוי, נלקח מדגם מקרי של 30 תאי שטח מכל אחד משימושי הקרקע הללו (סה"כ 330 תאי שטח).

המיפוי של 30% מהתאים של שימושי הקרקע הבנויים עליהם נעשו הבקורות נבדק על ידי בחינת חזיתות המבנים באמצעות Google Street View (גוגל תצוגת רחוב).

מקור נוסף ששימש לאימות הנתונים היה האורתופוטו משנת 2014. עבור 46% מתאי השטח שנדגמו נבחן המיפוי באמצעות מקור זה. כמו כן, היו תאי שטח בהם אומת המיפוי באמצעות בירורים טלפוניים או באמצעות חיפוש באינטרנט (24% מהתאים הנ"ל).

נמצא כי הדיוק הכולל של מיפוי שימושי הקרקע הבנויים, שנבחרו לצורך הבקרה הנו 71.52%, הדיוק הכולל המשוקלל הוא 74% וטעות הדגימה של הדיוק הכולל המשוקלל היא 3.43%.

ערך אינדקס Kappa שהתקבל הוא 0.69. חושב אינדקס Kappa משוקלל במטרה להציג את Kappa במונחים של מדגם פרופורציוני. התקבל ערך של 0.7 עבור אינדקס Kappa משוקלל.

נוכל להגיד, באופן שמרני למדי, כי הערך האמיתי של Kappa הוא כ-0.7. זהו ערך גבוה המעיד על הצלחה של הניבוי דהיינו התאמה/הסכמה גבוהה בין המיפוי בשכבה ובין מקורות האימות.

נבדקה גם הצלחת המיפוי של שימושי הקרקע הללו ביחס לאימות: נמצא כי עבור שימושי הקרקע "שירותי דת", "חירום והצלה", "בריאות ורווחה" ו"שירותי מינהל ציבורי" אמינות המיפוי נמוכה – 60% ומטה. עם זאת, האחוז של כל אחד משימושי הקרקע הללו מתוך תאי השטח, שמופו בשכבה לפי שימושי הקרקע הבנויים שנבחרו, נמוך⁴⁵.

⁴⁵ הנתון המדויק מצוי בטבלה 7.

עבור שימושי הקרקע "חינוך והשכלה", "מסחר ומשרדים" ו"תעשייה" אמינות המיפוי בינונית ועבור שימושי הקרקע "תרבות פנאי וספורט", "מלונאות ואירוח כפרי", "מסחר ומגורים במרכז עירוני" ו"בתי עלמין" אמינות המיפוי גבוהה – מעל 80%.

7.2 התייחסות לשאלות שהועלו בראשיתה של העבודה

בראשיתה של העבודה הועלו מספר סוגיות מרכזיות שיש לבחון בנוגע למיפוי:

7.2.1 לשאלה "עבור כמה אחוזים מהשטח שמופה יש התאמה בין המיפוי למציאות?" ניתן בעבודה מענה באמצעות המדדים: "הדיוק הכולל" (Overall Accuracy), "דיוק כולל משוקלל", אינדקס Kappa ו-"אינדקס Kappa משוקלל" (Weighted Kappa). לשם בחינת הדיוק לפי קטגוריה ניתנה לשאלה זו מענה באמצעות המדדים: "דיוק צרכן" (User's Accuracy) ו-"אינדקס kappa עבור שימוש קרקע נתון" - Conditional Kappa (User's).⁴⁶

7.2.2 לשאלה "מהו הדפוס של הטעויות? כלומר: איזה קטגוריות לא מופו נכון ומהן הקטגוריות לפיהן הן מופו בטעות" ניתן מענה במטריצות העירוב.

7.2.3 לשאלה השלישית "מה מידת החומרה שמייחס המשתמש במפה לטעויות שונות?" לא היתה כוונה לתת מענה בעבודה זו. הסיבה היא ששכבת שימושי קרקע 2014 נועדה לשרת מגוון משתמשים. כל אחד מהם זקוק למידע על שימושי הקרקע למטרות אחרות והוא מדרג את חומרת הטעויות בהתאם למטרות השימוש שלו.

7.3 מסקנות

ניתן ליישם את השיטה שננקטה גם בעבודות נוספות. יתרונותיה של השיטה:

7.3.1 היא מאפשרת יכולת ניבוי גם עבור קטגוריות קטנות, שאין מהן הרבה תאים בשכבה (בניגוד למדגם פרופורציוני) וגם עבור קטגוריות גדולות.

7.3.2 על ידי שינוי הנוסחה של אינדקס Kappa (Kappa משוקלל לפי האחוז של הקטגוריה מהסה"כ) השיטה מאפשרת להציג את ה-Kappa, שהיה מתקבל, במונחים של מדגם פרופורציוני.

⁴⁶ חושב גם המדד "דיוק יצרן" (Producer's Accuracy) אך הוא לא עונה על השאלה הזו אלא על השאלה "מה הסיכוי שתא שטח בנתוני האימות מקטגוריה נתונה מופה נכון".

רשימת מקורות

- אברהם, א' (בהכנה). יצירת בסיס נתונים ארצי עדכני בנושא שימושי הקרקע באמצעות שילוב בין מידע ממרשמים מנהליים ומשכבות מידע גאוגרפיות ובין מידע המופק באמצעות חישה מרחוק. עבודה למסלול הקידום המקצועי. ירושלים: הלשכה המרכזית לסטטיסטיקה.
- גוגל (אין תאריך א). [גוגל תצוגת רחוב (Google Street View) תמונות ב-360 מעלות ברמת הרחוב]. אוחר ב 4 בספטמבר 2016 מתוך <https://www.google.co.il/maps/@31.7963186,35.175359,12z>
- גוגל (אין תאריך ב). [גוגל ארת' (Google Earth) תצלומי לוויין של כדור הארץ ברזולוציית צילום גבוהה]. אוחר ב 4 בספטמבר 2016 מתוך <https://www.google.co.il/maps/@31.7963186,35.175359,29948m/data=!3m1!1e3>
- גוגל תצוגת רחוב (אין תאריך). ויקיפדיה. אוחר מתוך https://he.wikipedia.org/wiki/%D7%92%D7%95%D7%92%D7%9C_%D7%AA%D7%A6%D7%95%D7%92%D7%AA_%D7%A8%D7%97%D7%95%D7%91
- הלשכה המרכזית לסטטיסטיקה (2016). דמוגרפיה של עסקים: מקבץ נתונים סטטיסטיים ממרשם העסקים 2011–2014. פרסום מס' 1636. אוחר מתוך http://www.cbs.gov.il/publications16/1636_business_demography_2014/pdf/print.pdf
- הוועדה הבין-משרדית למערכות מידע גאוגרפיות (2014). תקן מערך קודים אחיד למיפוי תכסיות ושימושי קרקע. אוחר מתוך http://mapi.gov.il/gisForum/tzevet_technology/1.pdf
- לוגסי, ר', לוין, נ', רמון, א' ובן דור, א' (2004). שימוש בחישה מרחוק ככלי לאיתור חממות בנף החקלאי בישראל, דו"ח סופי, מוגש לנקודת ח"ן – הקרן לקידום ערכי נופ וסביבה באזורים חקלאיים בישראל.
- לוין, נ' (2003). מושגי יסוד ב-GIS וחישה מרחוק. אוחר מתוך <http://geography.huji.ac.il/personal/Noam%20Levin/2003Levin%20-GIS-course-theory.pdf>
- מזור, ש', סבוראי, ט' ובר (קוטיאל), פ' (2005). מודל מבוסס מ"ג לחיזוי כיסוי הצומח המעוצה בשטחים מקוטעים טבעיים בנגב הצפוני. יער, 7, 26–34. אוחר מתוך

http://www.kkl.org.il/files/HEBREW_PICTURES/gilayon_yaar/model_pdf.pdf

מערכת מידע גאוגרפית (אין תאריך). ויקיפדיה. אוצר מתוך

<https://he.wikipedia.org/wiki/%D7%A2%D7%A8%D7%9B%D7%AA%D7%9E%D7%99%D7%93%D7%A2%D7%92%D7%90%D7%95%D7%92%D7%A8%D7%A4%D7%99%D7%AA>

ובח, מ' (2013). *מרשם דירות*. מצגת. ירושלים: הלשכה המרכזית לסטטיסטיקה.

Adam, E., Mutanga, O., Odindi, J., & Abdel-Rahman, E. M. (2014). Land-Use/Cover Classification in a Heterogeneous Coastal Landscape Using RapidEye Imagery: Evaluating the Performance of Random Forest and Support Vector Machines Classifiers. *International Journal of Remote Sensing*, 35(10), 3440-3458.

doi: 10.1080/01431161.2014.903435

Ballestores, F., & Qui, Z. (2012). An Integrated Parcel-Based Land Use Change Model Using Cellular Automata and Decision Tree. *Proceedings of the International Academy of Ecology and Environmental Sciences*, 2(2), 53–69. Retrieved from

[http://www.iaees.org/publications/journals/piaees/articles/2012-2\(2\)/an-integrated-parcel-based-land-use-change-model.pdf](http://www.iaees.org/publications/journals/piaees/articles/2012-2(2)/an-integrated-parcel-based-land-use-change-model.pdf)

Berghout, M., Robinson, H., & Owen, G. (2008). *Revised Land Use Mapping of the Northern Territory 2008*. Technical Report No. 20/200D. Department of Natural Resources, Environment, the Arts and Sport, Palmerston, Northern Territory. Retrieved from http://www.agriculture.gov.au/abares/aclump/Documents/LUMP08_Rpt_NT_ALUMv6.pdf

Congalton, R. G. (1991). A Review of Assessing the Accuracy of Classification of Remotely Sensed Data. *Remote Sensing of Environment*, 37, 35–46. Retrieved from <http://uwf.edu/zhu/evr6930/2.pdf>

FAO (2005). *Land Cover Classification System: Classification Concepts and User Manual*. Retrieved from http://www.glcn.org/downloads/pub/docs/manuals/lccs/LCCS2-manual_en.pdf

- Jiang, S., & Liu, D. (2011). On Chance-Adjusted Measures for Accuracy Assessment in Remote Sensing Image Classification. In *ASPRS 2011 Annual Conference*. Retrieved from <http://www.asprs.org/wp-content/uploads/2010/12/Jiang.pdf>
- Liu, C., Frazier, P., & Kumar, L. (2007). Comparative Assessment of the Measures of Thematic Classification Accuracy. *Remote Sensing of Environment* 107, 606–616.
doi: 10.1016/j.rse.2006.10.010
- Lowry, J. H., Ramsey, R. D., Boykin, K., Bradford, D., Comer, P., Falzarano, S., ... Wolk, B. (2005). *Southwest Regional Gap Analysis Project: Final report on Land Cover Mapping Methods*. Retrieved from http://earth.gis.usu.edu/swgap/swregap_landcover_report.pdf
- Manandhar, R., Odeh, I. O. A., & Ancev, T. (2009). Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data Using Post-Classification Enhancement. *Remote Sensing*, 1(3), 330-344.
doi:10.3390/rs1030330
- Rossiter, D. G. (2004). *Technical Note: Statistical Methods for Accuracy Assessment of Classified Thematic Maps*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.247&rep=rep1&type=pdf>
- Rozenstein, O., & Karnieli, A. (2011). Comparison of Methods for Land-Use Classification Incorporating Remote Sensing and GIS Inputs. *EARSeL eProceedings*, 10, (1), 27–45. Retrieved from http://www.e proceedings.org/static/vol10_1/10_1_rozenstein1.pdf
- RS/GIS Laboratories Utah State University (2003). *Kappa Tool User's Guide*. Retrieved from <http://www.gis.usu.edu/~chrisg/avext/downloads/kappa.pdf>
- Stehman, S. V., & Czaplewski, R. L. (1998). Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. *Remote Sensing of Environment*, (64), 331–344. Retrieved from http://www.fs.fed.us/rm/pubs_other/rmrs_1998_stehman_s001.pdf
- Strahler, A., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., ... Woodcock, C. E. (2006). *Global Land Cover Validation*:

Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps. Retrieved from

http://www.nofc.forestry.ca/gofc-gold/Report%20Series/GOLD_25.pdf

Viera, A. J., & Garrett, J. M. (2005). Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine*, 37(5), 360–363. Retrieved from http://www1.cs.columbia.edu/~julia/courses/CS6998/Interrater_agreement.Kappa_statistic.pdf

פרויקט שימושי קרקע 2014

בשנת 2014 יוצר בתחום ממ"ג-גאוגרפיה בלשכה המרכזית לסטטיסטיקה בסיס נתונים מרחבי בנושא שימושי קרקע.

יוצרה רשת תאים ארצית רציפה בגודל 100*100 מטרים עם הערך של שימוש הקרקע העיקרי בכל תא שטח. קיימים בשכבה 37 שימושי קרקע וכיסויי קרקע. מטרת הפרויקט: השלמת חסרים בסטטיסטיקה הרשמית של מדינת ישראל בנושא קרקע, על ידי יצירת סדרה עיתית סטטיסטית בנושא שימושי הקרקע. הפרויקט מספק מידע עדכני ורציף על שימושי הקרקע במדינה⁴⁷ וכן על שימושי הקרקע בעריות ובמועצות המקומיות ביהודה והשומרון⁴⁸ ומאפשר הפקת נתונים עבור כל תיחום גאוגרפי.

שימושי הקרקע מופו בהתאם לתקן מערך קודים ארצי אחיד למיפוי תכסיות ושימושי קרקע,⁴⁹ שהוכן במסגרת הוועדה הבין-משרדית למערכות מידע גאוגרפיות, בהתאם לתקן של האו"ם לסיווג תכסית⁵⁰.

המידע מעודכן לסוף שנת 2013.

מקורות המידע: מידע ממרשמים מנהליים מעוגנים (מרשם דירות ומבנים, מרשם עסקים ומרשם התושבים), שכבות מידע גאוגרפיות, שהתקבלו ממשרדים ממשלתיים, מארגונים לא ממשלתיים ומגופים פרטיים ומידע מאורתופוטו לשם טיוב שכבת היערות. שיטת העבודה:

א. סיווג שכבת המבנים הארצית לפי שימושי קרקע, על סמך המרשמים המנהליים המעוגנים ועל סמך מגוון שכבות מידע גאוגרפיות.

ב. יצירת שכבה עבור כל אחד מ-37 שימושי הקרקע על סמך שכבת המבנים המסווגת בתוספת מעטפת ברוחב 20 מטרים ועל סמך שכבות מידע גאוגרפיות נוספות.

ג. שילוב השכבות לשכבה אחת לפי סדר היררכי, שנקבע בין שימושי הקרקע. הסדר ההיררכי הוא סדר הקודים בשכבת שימושי הקרקע, כאשר קוד מספר אחד הוא במקום הגבוה ביותר בהיררכיה.

⁴⁷ הפרויקט אינו ממפה מחנות צבא ולא שטחים ביטחוניים.

⁴⁸ כאמור, בשטחי יהודה והשומרון ניתנים נתונים על שימושי הקרקע רק בתחום השטח הבנוי של עריות ומועצות מקומיות. לא מופו המועצות האזוריות בשטחי יהודה והשומרון. לפיכך לדוגמא לא מוצגים שטחים בבקעת הירדן.

⁴⁹

- ד. טיוב שכבת היערות באמצעות מיון מבוקר של אורתופוטו בשיטה מתחום חישה מרחוק.
- ה. התמרה של השכבה הוקטורית שנוצרה לשכבה רסטרית 0.5 מטרים לפיקסל.
- ו. יצירת רשת תאים וקטורית (fishnet) בגודל תא 100 על 100 מטרים.
- ז. מתן ערך של שימוש הקרקע העיקרי לפי השכבה הרסטרית.

התוצר הסופי כולל שכבה רסטרית של תאים ושכבה וקטורית נקודתית.

התוצר מכיל שתי שכבות:

א. שכבה רסטרית ארצית - גודל פיקסל 100 מטרים. עבור כל פיקסל קיים נתון על קוד שימוש הקרקע לפי הלשכה לסטטיסטיקה, שמו של שימוש הקרקע וקוד שימוש הקרקע התואם לפי תקן מערך הקודים הארצי.

שם השכבה: land_use2014g.img

גודל תא 100 על 100 מטרים.

פורמט Erdas imagine

השכבה נתונה ברשת ישראל החדשה בהיטל GRS_1980_Transverse_Mercator

תיאור השדות בשכבה:

שם השדה	תיאור השדה	סוג השדה	רוחב השדה
Value	קוד שימוש הקרקע שניתן לשכבה בלשכה לסטטיסטיקה	long	9
Landuse_co	קוד שימוש הקרקע שניתן לשכבה בלשכה המרכזית לסטטיסטיקה	double	16
Landuse_na	שם שימוש הקרקע	string	254
Teken_code	קוד שימוש הקרקע התואם לפי התקן הארצי האחד למיפוי תכסיות ושימושי קרקע	double	16

לשכבה מצורפים גם הפריטים הבאים:

א. 1 טבלה עם תקן סימבולוגיה לשימושי הקרקע.

שם הטבלה: "key_landuse_teken_symbology".

בטבלה נתונים ערכי צבעים rgb לפי שלוש רמות פירוט:

ערכי rgb עבור כל אחד מ-37 שימושי הקרקע.

ערכי rgb עבור 26 קטגוריות עיקריות של שימושי קרקע (השדה "קטגוריה כללית-עירוני").

ערכי rgb עבור 9 קטגוריות עיקריות של שימושי קרקע (השדה "קטגוריה כללית-ארצי").

א.2. בהתאם לערכים בטבלה א.1 מצורפים גם שלושה קובצי lyr (לצביעת השכבות) לפי הפירוט הבא:

א.2.1. "ראסטר_37_שימושי הקרקע 251115" – לצביעה של כל אחד משימושי הקרקע.

א.2.2. "ראסטר_37_שימושי קרקע_26_קטגוריות עיקריות_251115" – לצביעה לפי 26 קטגוריות עיקריות של שימושי קרקע.

א.2.3. "ראסטר_37_שימושי קרקע_9_קטגוריות עיקריות_251115" – לצביעה לפי 9 קטגוריות עיקריות של שימושי קרקע.

הערה: שימושי קרקע, שאוחדו לקטגוריה משותפת בחלוקה ל-26 קטגוריות, יקבלו אותו צבע גם בצביעה של שימושי הקרקע בנפרד. ההבדל בין א.2.1 לבין א.2.2 הוא שב-א.2.1 מופיעה הרשימה המלאה של שימושי הקרקע.

כדי להשתמש באחד מקובצי ה-lyr על מנת לצבוע את השכבה הרסטריית יש לבצע את הפעולות הבאות: להיכנס לכרטיסיה symbology בחלון ה-properties של השכבה ולצבוע את השכבה לפי ערך ייחודי (unique_values). אפשר לצבוע לפי כל אחד מהשדות בשכבה. לדוגמה לפי השדה value או לפי השדה landuse_co ואז ללחוץ על "import" ולייבא את קובץ ה-lyr לפיו רוצים לצבוע.

ב. שכבה וקטורית נקודתית – שכבה המכילה את נקודות המרכז של כל אחד מתאי השטח משכבה א'.

בשכבה מידע על מספר חד-ערכי לכל אחד מתאי השטח, קואורדינטות y, x של מרכז כל תא שטח, קוד שימוש הקרקע לפי הלשכה לסטטיסטיקה, שמו של שימוש הקרקע וקוד שימוש הקרקע התואם לפי תקן מערך הקודים הארצי.

שם השכבה: land_use2014_points

סוג השכבה: נקודתית

פורמט: Shape File

השכבה נתונה ברשת ישראל החדשה בהיטל GRS_1980_Transverse_Mercator

תיאור השדות בשכבה:

שם השדה	תיאור השדה	סוג השדה	רוחב השדה
TA_NUM	מספר חד ערכי ייחודי לכל אחד מתאי הרשת	long	9
X_COORD	קואורדינטת x של הנקודה שמייצגת את מרכז תא השטח	double	19
Y_COORD	קואורדינטת y של הנקודה שמייצגת את מרכז תא השטח	double	19
LANDUSE_CO	קוד שימוש הקרקע שניתן לשכבה בלשכה המרכזית לסטטיסטיקה	short	4
LANDUSE_NA	שם שימוש הקרקע	text	60
TEKEN_CODE	קוד שימוש הקרקע התואם לפי התקן הארצי האחד למיפוי תכסיות ושימושי קרקע	double	19

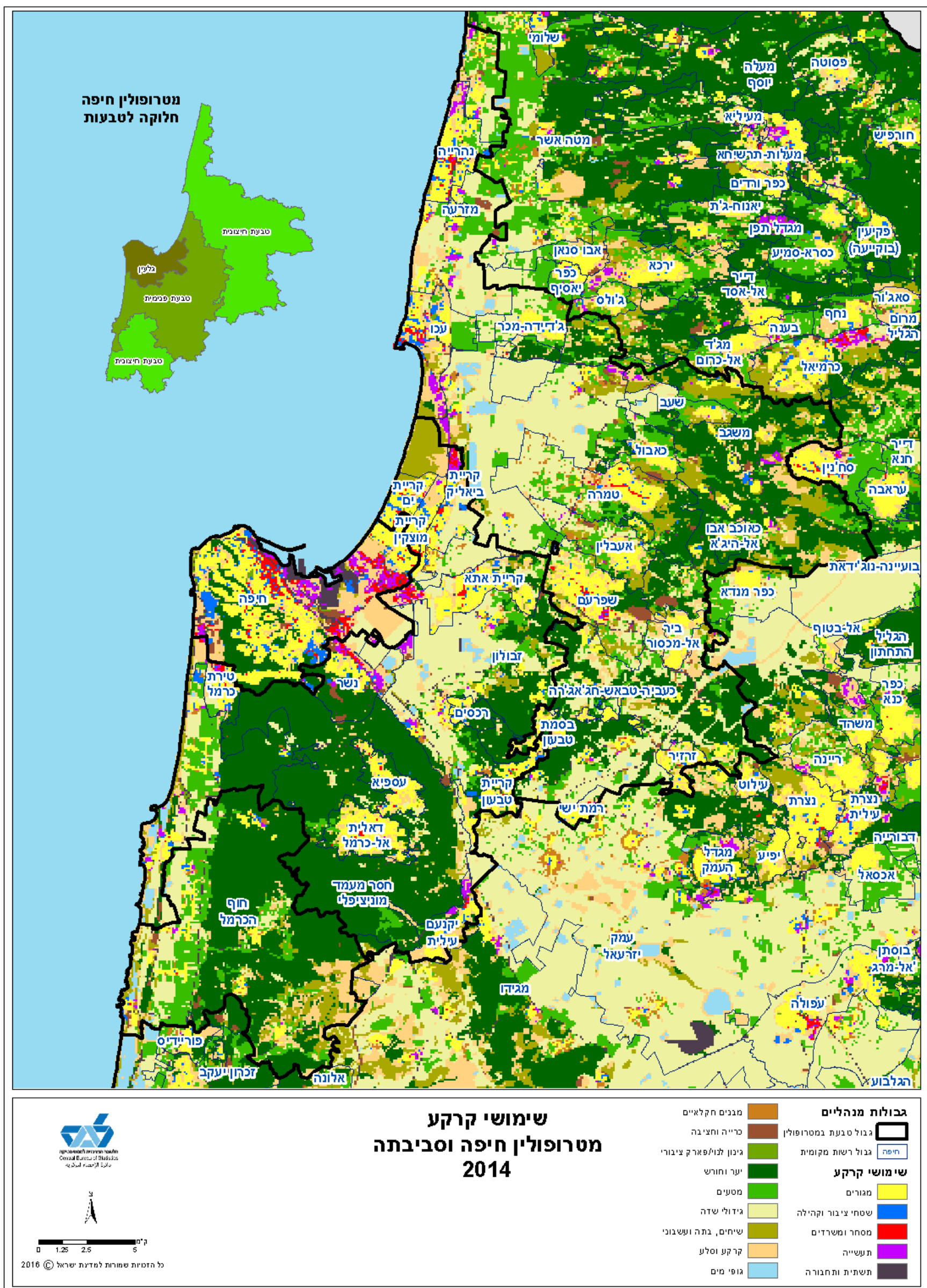
שימושי הקרקע שמופיעים בשתי השכבות לפי הסדר ההיררכי, לפיו שולבו בתהליך המיפוי:

TEKEN_CODE	LANDUSE_NAME	LANDUSE_CODE
311100	חינוך והשכלה	1
311200	בריאות ורווחה	2
311300	שירותי דת	3
311400	חירום והצלה	4
311500	שירותי מינהל ציבורי	5
311600	בתי עלמין	6
311700	תרבות פנאי וספורט	7
311800	מלונאות ואירוח כפרי	8
315100	מסחר ומגורים במרכז עירוני	9
321600	תחנות תדלוק ושירותי דרך	10
312000	מסחר ומשרדים	11
313000	תעשייה	12
321500	מכלולי תחבורה מסילתית	13
321400	מכלולי תחבורה אווירית	14
321300	מכלולי תחבורה ימית	15
322000	חשמל ותקשורת	16
323000	סילוק פסולת	17
324000	תשתיות רטובות סגורות	18
311000	שירות לציבור ללא אפיון	20
314000	מגורים	21
332000	מתחמי גידול בעלי חיים	22
331000	בתי צמיחה	23
330000	מבנים חקלאיים ללא אפיון	24
321200	חניונים ותחנות מרכזיות	25
341000	כרייה וחציבה	26
342000	חפירה ארכיאולוגית	27
343000	חפור מופר אחר	28
321100	כבישים	29
510000	גופי מים מלאכותיים	31
610000	גופי מים טבעיים	32
120000	גינון לנוי פארק ציבורי	33
210000	יער וחורש	34
111000	מטעים	35
112000	גידולי שדה	36
220000	שיחים בתי ועשבונים	37
410000	אזור סלעי	38
420000	קרקע	39

קוד מספר 1 מייצג את שימוש הקרקע הגבוה ביותר בהיררכיה.

לרכישה של התוצרים יש לפנות לתחום מידע ודוברות בטל' 02-6592666, 03-5681933,

בדוא"ל: info@cbs.gov.il



נספח 3: שימושי הקרקע – קטגוריות לצורך בקרת התוצר

קוד שימוש הקרקע בשכבה	שם שימוש הקרקע	מספר התאים שמופו בשכבה	קטגוריה
1	חינוך והשכלה	4,862	מתחמים בנויים
2	בריאות ורווחה	816	מתחמים בנויים
3	שירותי דת	350	מתחמים בנויים
4	חירום והצלה	144	מתחמים בנויים
5	שירותי מינהל ציבורי	485	מתחמים בנויים
6	בתי עלמין	875	מתחמים בנויים
7	תרבות פנאי וספורט	3,557	מתחמים בנויים
8	מלונאות ואירוח כפרי	385	מתחמים בנויים
9	מסחר ומגורים במרכז עירוני	1,139	מתחמים בנויים
10	תחנות תדלוק ושירותי דרך	551	מכלולי תשתית
11	מסחר ומשרדים	4,702	מתחמים בנויים
12	תעשייה	7,034	מתחמים בנויים
13	מכלולי תחבורה מסילתית	65	מכלולי תשתית
14	מכלולי תחבורה אווירית	2,190	מכלולי תשתית
15	מכלולי תחבורה ימית	695	מכלולי תשתית
16	חשמל ותקשורת	114	מכלולי תשתית
17	סילוק פסולת	332	מכלולי תשתית
18	תשתיות רטובות סגורות	199	מכלולי תשתית
20	שירות לציבור ללא אפיון	482	מתחמים בנויים
21	מגורים	83,030	מתחמים בנויים
22	מתחמי גידול בעלי חיים	3,756	מבנים חקלאיים
23	בתי צמיחה	9,394	מבנים חקלאיים
24	מבנים חקלאיים ללא אפיון	2,331	מבנים חקלאיים
25	חניונים ותחנות מרכזיות	492	מכלולי תשתית
26	כרייה וחציבה	3,607	חפור מופר
27	חפירה ארכיאולוגית	3	חפור מופר
28	חפור מופר אחר	13,469	חפור מופר
29	כבישים	13,423	מכלולי תשתית
31	גופי מים מלאכותיים	24,734	גופי מים
32	גופי מים טבעיים	32,450	גופי מים
33	גינון לנוי פארק ציבורי	2,231	גינון לנוי פארק ציבורי
34	יער וחורש	158,727	יער וחורש
35	מטעים	110,654	מטעים
36	גידולי שדה	316,810	גידולי שדה
37	שיחים בתה ועשבוני	118,626	שיחים בתה ועשבוני
38	אזור סלעי	592,064	אזור סלעי
39	קרקע	721,168	קרקע
	סך הכל	2,235,946	

נספח 4: שלבי היצירה של נוסחה (7) - אינדקס Kappa משוקלל (Weighted Kappa)

הגדרת הסימונים בנספח זה מופיעים בסעיף 6.1.2.

א. אינדקס Kappa לא משוקלל נוסחה (6) (RS/GIS Laboratories Utah State University,) (2003)

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})} \quad (6)$$

ב. שלבי היצירה של נוסחה מספר (7) - אינדקס Kappa משוקלל (Weighted Kappa)

ב.1 גזירת הנוסחה – הכפלת כל ערך בטבלה בשוליים שהיו צפויים להתקבל תחת דגימה פרופורציונית של תאי שטח

$$\begin{aligned} K_w &= \frac{N \sum_{i=1}^r x_{ii} * \frac{N}{ni} * W_i - \sum_{i=1}^r \left[\left(\sum_{j=1}^r x_{ij} * \frac{N}{ni} * W_i \right) \left(\sum_{k=1}^r x_{ki} * \frac{N}{ni} * W_k \right) \right]}{N^2 - \sum_{i=1}^r \left[\left(\sum_{j=1}^r x_{ij} * \frac{N}{ni} * W_i \right) \left(\sum_{k=1}^r x_{ki} * \frac{N}{ni} * W_k \right) \right]} \\ &= \frac{N^2 \sum_{i=1}^r x_{ii} \frac{W_i}{ni} - N^2 \sum_{i=1}^r \frac{W_i}{ni} \sum_{k=1}^r \frac{W_k}{ni} x_{ki} \sum_{j=1}^r x_{ij}}{N^2 - N^2 \sum_{i=1}^r \frac{W_i}{ni} \sum_{k=1}^r \frac{W_k}{ni} x_{ki} \sum_{j=1}^r x_{ij}} \\ &= N^2 \frac{\sum_{i=1}^r \frac{x_{ii} * W_i}{ni} - N^2 \sum_{i=1}^r \left(\sum_{j=1}^r \frac{x_{ij} * W_i}{ni} \right) \left(\sum_{k=1}^r \frac{x_{ki} * W_k}{ni} \right)}{N^2 - N^2 \sum_{i=1}^r \left(\sum_{j=1}^r \frac{x_{ij} * W_i}{ni} \right) \left(\sum_{k=1}^r \frac{x_{ki} * W_k}{ni} \right)} \end{aligned}$$

ב.1 רישום אלטרנטיבי -

נגדיר: $\tilde{X}_{ij} = W_i * x_{ij}$ ו- $\tilde{X}_{ii} = W_i * x_{ii}$

אזי $\tilde{X}_{+i} = \sum_{k=1}^r W_k * x_{ki}$ ו- $\tilde{X}_{i+} = \sum_{j=1}^r \tilde{X}_{ij} = \sum_{j=1}^r W_i * x_{ij}$

ומתקבל

$$K_w = \frac{N^2 \sum_{i=1}^r \frac{\tilde{X}_{ii}}{ni} - N^2 \sum_{i=1}^r \frac{\tilde{X}_{i+}}{ni} * \frac{\tilde{X}_{+i}}{ni}}{N^2 - N^2 \sum_{i=1}^r \frac{\tilde{X}_{i+}}{ni} * \frac{\tilde{X}_{+i}}{ni}}$$

Abstract

The Central Bureau of Statistics (CBS) has released new official data on land use in Israel at the end of 2013.

The database used for the information on land use was prepared in 2014 through the geographic Information System (GIS) of the CBS. The information is based on processing and integration of a variety of administrative sources that were received from government ministries and nongovernment and private organizations. Data source includes geocoded registers: Buildings and Dwelling Register, Business Register and Central Population Register, GIS layers that were received from: Central Bureau of Statistics, Survey of Israel - the government agency for Mapping, Geodesy, Cadastre and Geoinformatics, Jewish National Fund, Mapa - Mapping & Publishing Ltd., Ministry of Agriculture & Rural Development, Ministry of National Infrastructures, Energy and Water Resources, Ministry of Education, Magen David Adom in Israel - the National Rescue Organization, and Israel Fire Rescue Services. Orthophoto was used to improve the forest layer.

A continuous nationwide network of cell sizes measuring 100x100 meters was created and a value provided for each cell area. The database includes 37 land use types.

Land use was mapped according to the national standard system of symbols for mapping land cover and land use. The national system was prepared by an inter-ministerial committee for geographic information systems, in accordance with the United Nations system.

Methodology: The national building layer was classified by land use due to administrative registers and a variety of GIS layers. A geographic layer was prepared for each of 37 land uses by placing 20 meters buffer around the buildings and gathering additional layers. These 37 GIS layers were union into one land use layer by hierarchical order. The layer was converted into raster (0.5 meter pixel). A continuous nationwide network of cell sizes measuring 100x100 meters was created and a value provided for each cell area due to the majority value of land use. The final product is a raster format.

Land use information is crucial for planning and policy making in many fields. Therefore, there is a need to evaluate the data and to know the reliability of the mapping results.

The objective of this paper is to assess the accuracy of the geographic database on land usage, and to what extent there is a match between the layer mapping land usage between 2014 as well as information from other sources (orthophoto and Google Street View) which were used to validate the results of the mapping. The sampling unit of 100x100 meters is used because land use information is given for this cell size.

In the first stage, a fixed size sample was taken from each category, regardless of the category size, to allow for predictability of even smaller categories of land use. In the second stage, a method was developed for the correct calculation of the Kappa index of agreement. The correction of the formula enables the Kappa to be viewed in terms of proportional sampling. Each category is weighted according to the percentage of the total.

The Land use is grouped into twelve categories. It was found that the overall accuracy of the mapping category is 82.75%. The overall weighted accurateness is 80%. The Cohen's Kappa index is 0.81, and the weighted kappa as calculated in this study is 0.75.

Most of the categories achieved over 85% user's accuracy with the exception of "dig areas" (e.g. quarries, archeological excavation) and "cultivated fields" which had only about 60%.

Also, we assessed eleven land uses in built up areas. The overall accuracy for the mapping is 71.52%. The weighted overall accuracy is 74%. The Cohen's Kappa index is 0.69, and the weighted kappa as calculated in this study is 0.7. "Culture, leisure and sport," "Hotels and rural hosting," "Commerce and residence in the main streets" and "Cemeteries" achieved over 80% user's accuracy. "Education," "Commerce and offices" and "Industry" produced between 67% to 77% user's accuracy. For "religious services," "Emergency and rescue," "Health and welfare" and "Public administration services" user's accuracy is 60% or less.

Keywords: land use, land cover, accuracy assessment, validation

I would like to thank Dr. Noam Cohen, Senior Director of Sector in the Senior Department of Statistical Methodology for statistical consultation about the evaluation of the Land Use Database and for his professional guidance.

I would like to thank Eyal Maharian, Director of Sector-Geography Information System who has chosen me for planning and preparing the Land Use Database. I Also thank him for his helpful suggestions and for his professional knowledge.

I would like to thank Dr. Rakefet Shafran-Natan from Ben-Gurion University of the Negev who taught me how to use methods for land use classification from Remote-sensing data.

Thanks to Polina Blechman from Geography Information System Sector who joined me for improving the forest layer.

Thanks to Oren Raz from Geography Information System Sector who prepared the map of Land use in Haifa metropolitan area according to the land use layer which I have prepared.

I would like to thank Shahar Katz, ,Senior Director of Senior Department of Information Technology for his helpful comments.

The Central Bureau of Statistics (CBS) encourages research based on CBS data, such as this work. Works of research of this sort are not official publications of the CBS, and therefore the opinions and conclusions expressed in these publications are those of the authors and do not necessarily represent those of the CBS.

Published by the Central Bureau of Statistics, 66 Kanfei Nesharim St.,
Corner Bachi St., P.O.B 34525, Jerusalem 91342, Israel
Tel. 972-2-6592666; Fax: 972-2-6521340
Internet Site: www.cbs.gov.il
E-Mail: info@cbs.gov.il

WORKING PAPER SERIES

No. 100

Evaluation and Accuracy Assessment of the CBS Land Use Map

Efrat Avraham

September 2016