



המארג
התכנית הלאומית
להערכת מצב הטבע

נספח טכני ניטור זוחלים

דו"ח מצב הטבע
ישראל 2023 | המגוון הביולוגי

ניטור זוחלים

תוכן עניינים

1.	כללי	2
2.	עיבוד וסינון לפני ניתוח	2
3.	ניתוח	4
4.	מקורות	6

1. כללי

כל שלבי הניתוח, העיבוד וסינון הנתונים בוצעו בסביבת R (R Core Team, 2023) גרסה 4.3.0. כל השלבים תועדו בקובץ R Markdown:

Analyzing the reptile data for State of Nature Report 2023.Rmd

פלט הקובץ הזה מופיע בתור קובץ html:

Analyzing-the-reptile-data-for-State-of-Nature-Report-2023.html

הגרפים שורטטו באמצעות החבילות ggplot2 (Wickham, 2016), גרסה 3.4.4 ו-Cairo (Urbanek & Horner, 2022) גרסה 1.6-2.

2. עיבוד וסינון לפני ניתוח

- 1. פרוטוקול הניטור** נבדל בין היחידות החוליות (חולות מישור החוף וחולות מערב הנגב) לבין היחידות האחרות. ביחידות החוליות הוטמנו מלכודות נפילה לצד שביל טשטוש, בנוסף לסריקה רגלית, בעוד שביתר היחידות הפרוטוקול כלל סריקה רגלית בלבד. מכאן שמאמץ הדיגום ביחידות החוליות היה גבוה יותר מאשר ביתר היחידות, ולא ניתן להשוות במישרין את התוצאות מבלי לקחת זאת בחשבון. על כן, הניתוח התמקד ברמת היחידה ולא השווה בין יחידות הניטור השונות.
- 2. מסד הנתונים** כלל תצפיות בזוחלים בסקרי זוחלים בכל יחידה, וכלל את מחזור הניטור, האתר, החלקה, קואורדינטות (נ.צ.), משתני הניטור בהתאם ליחידה (קרוב/רחוק מישובים או מחקלאות, בית גידול במישורי הלס ואופי הדיונה ביחידות החוליות, תת יחידה ביער המחטני הנטוע),

תאריך, שעת התחלה, מין הזוחל שנצפה, מספר הפרטים, פעילות הזוחל בעת התצפית, אורך הגוף ואורך הזנב (במידה ונלכד), וקבוצת גיל (צעיר, מתבגר, תת-בוגר, בוגר או ללא מידע). דיגומים בהם לא נצפו זוחלים כלל הוספו למסד הנתונים עם הפרטים הרלוונטיים.

3. נתונים משנת הפיילוט (2013) ביחידת הלס הושמטו ממסד הנתונים משום שפרוטוקול הניטור היה שונה אז לעומת המחזורים המאוחרים יותר.

4. כאשר **הקואורדינטות** של אותה החלקה נבדלו בין מחזורים (4 חלקות בסה"כ), השתמשנו בקואורדינטות הממוצעות. המרחק בין הקואורדינטות השונות של אותה החלקה הגיע לכדי שתי אלפיות המעלה (כ-200-250 מ') לכל היותר. עבור שתי חלקות (להבים קרוב 1 ולהבים קרוב 6) לא תועדו קואורדינטות כלל.

5. תוקנו שגיאות הקלדה **בשמות האתרים והחלקות**. שמות האתרים והחלקות גם הואחדו. למשל, האתר "חצרים" במחזור הניטור הראשון (2014) ביחידת הלס נקרא במחזורים מאוחרים יותר "סיירת שקד", אז שם האתר שונה רטרואקטיבית ל"סיירת שקד" גם במחזור הניטור הראשון. בנוסף, כאשר שמות החלקות היו חסרים, נקבע שם חלקה חדש כהלחם של שם האתר, משתנה הניטור המרחבי הרלוונטי ליחידה (למשל, קרוב/רחוק מישוב) והקואורדינטות.

6. לאחר קביעת הקואורדינטות ושמות החלקות, נבחנה האפשרות שניתנו שמות שונים לאותן החלקות עצמן במחזורים שונים ע"י בחינת כלל המרחקים מכל החלקות השייכות לאותם אתרים לאורך כלל מחזורי הניטור. כאשר נמצאו שתי חלקות שונות באותו אתר בעלות אותם מאפיינים סביבתיים (למשל, שתיהן קרובות לישוב / חקלאות) ומרחקן פחות מ-5 מ' ([טווח השגיאה של מכשיר GPS](#)) זו מזו, הן אוחדו לכדי חלקה אחת, ששמה ומיקומה נקבעו שרירותית לפי אחת החלקות שהרכיבו אותה.

7. ב-20 דיגומים (כולם בחולות מישור החוף במחזור הניטור הראשון, בשנת 2014) דווחו שני **תאריכי ניטור** שונים אך עוקבים (לדוגמה, 30 בספטמבר 2014 ו-1 באוקטובר 2014) עבור אותו דיגום (שילוב של מחזור ניטור וחלקה). המקור של מצב כזה הוא פרוטוקול הניטור, בו מציבים מלכודות נפילה בשעות אחר הצהריים ואוספים אותן בבוקר למחרת. במצב כזה, השתמשנו בתאריך המוקדם מבין השניים. תאריך הניטור שימש לשקלול העונתיות בניתוח (למשל, האם זוחלים פעילים יותר בקיץ לעומת החורף; ראו להלן).

8. **שעת התחלת הניטור** נקבעה בתור השעה המוקדמת ביותר באותו דיגום (שילוב של תאריך וחלקה). שעת התחלת הניטור שימשה לשקלול דגם הפעילות היממית בניתוח (למשל, האם זוחלים פעילים יותר או פחות בשעות הבוקר לעומת שעות אחר הצהריים; ראו להלן). שעת התחלת הניטור שונתה לשעה כפי שהייתה אילו הופעל שעון חורף (לפיו חצות היום חלות ב-11:40) כל ימות השנה, כדי להימנע מהסטה מלאכותית של מחוגי השעון בשל הפעלת שעון קיץ.

9. תאריך הדיגום שימש לצורך חישוב שני משתנים ששיקפו את העונתיות: סינוס וקוסינוס של המרחק ברדיאנים בין התאריך לבין היום הארוך בשנה (21 ביוני). משתנים אלו משמשים עבור גרסיה מחזורית (Flurry & Levri 1999; Lusa & Ahlin, 2020) המאפשרת שיא ושפל במהלך המחזור השנתי (באמצעות משתנים המתייחסים למרחק ברדיאנים מהיום הארוך בשנה). בדומה לכך, סינוס וקוסינוס המרחק ברדיאנים בין שעת התחלת הדיגום לבין השעה 12:00 (בשעון חורף כאמור) שימשו לשיקוף הדגם היממתי.
10. **טקסונומיה:** שמות המינים המדעיים והעבריים נקבעו לפי טבלת תכונות מיני הזוחלים (בהנחיית פרופ' שי מאירי). זוחלים שדווחו כתתי מינים (כגון זיקית ים תיכונית) הוגדרו לרמת המין (כגון זיקית).
11. **השפע** של כל מין נקבע כמספר הפרטים שנצפו בו בדיגום. סימני שדה (עקבות, ביצים, שברי ביצים ונשלים) לנוכחות המין נספרו כפרט בודד, ללא קשר למספרם בפועל. דיווחים על מעל תשעה פרטים בתצפית אחת (היו שלושה דיווחים כאלה בכל מסד הנתונים) נחשבו לשגיאות הקלדה והומרו לתצפית בפרט בודד. הנתונים נותחו ברמת השפע לדיגום (סקר בחלקה במחזור ניטור).
12. **השפע הכולל** הוגדר בתור סך השפע מכלל המינים שנצפו בדיגום.
13. **עושר המינים** הוגדר כמספר המינים (ללא התייחסות לתתי מינים שונים) שנצפו בדיגום. הנתונים נותחו ברמת הדיגום (סקר בחלקה במחזור ניטור).

3. ניתוח

1. **המשתנים המסבירים** של כל דיגום הוגדרו בתור המשתנים שהוגדרו מראש עבור יחידת הניטור (קרוב / רחוק מישובים או מחקלאות, בית גידול ביחידת מישורי הלס, סוג הדיונה ביחידות החוליות, תת יחידה עבור היער המחטני הנטוע – להלן משתני הניטור), המגמה העתית (מספר השנים שחלפו מאז שנת תחילת הניטור – עבור השפע הכולל בלבד), לצד העונתיות (בכל היחידות) והדגם היממתי (עבור היחידות הלא חוליות בלבד, שכן ביחידות החוליות התצפיות הן במלכודות הפועלות במשך שעות רבות). בנוסף, ביחידת חולות מישור החוף הוגדר גם האזור - צפון (קיסריה בלבד) או דרום (אשדוד, אשקלון או ניצנים). ביחידות האחרות, זהות האתר הוגדרה כמשתנה מסביר נוסף – שלא הושמט בהליך ברירת המודלים גם במידה והפרוצדורה רמזה שיש להשמיט אותו, וזאת משום שהוא מייצג את המבנה המרחבי של הנתונים.
2. ראשית, שורטטו גרפים המתארים כיצד מדד המגוון הביולוגי משתנה לפי משתני הניטור הרלוונטיים ליחידה ובין מחזורי הניטור. גרפים אלו נועדו כעזר חזותי לניתוח הנתונים על מנת להבין מה מהמשתנים המסבירים צפוי להראות השפעה מובהקת סטטיסטית ולאיזה כיוון, אך לא כתחליף לניתוח סטטיסטי (ראו להלן).

3. עבור מדדי המגוון הביולוגי (עושר המינים או השפע הכולל), הותאם מודל לינארי מוכלל מעורב (Generalized Linear Mixed Model), בו מדד המגוון הביולוגי שנבחן היה המשתנה המגיב באמצעות החבילה `glmmTMB` (Brooks et al., 2017) גרסה 1.1.8 בסביבת R. סף המובהקות (אלפא) שנבחר היה 5%. בשלב הראשון, נבחרה משפחת ההתפלגות של המודל, באמצעות בחינת הערך ϕ_i , המוגדר בתור ה-deviance חלקי דרגות החופש של המודל. כאשר ערך זה היה קטן מ-1, הועדפה התפלגות פואסונית, וכאשר הוא היה גבוה מ-1, הועדפה התפלגות בינומית שלילית.
4. לאחר מכן, הותאם המודל המלא, בו מדד המגוון הביולוגי היה המשתנה המגיב וכלל את כל המשתנים המסבירים (ראו לעיל), וכן אינטראקציות בין משתני הניטור. לאחר התאמת המודל המלא כאמור, נבחנה השפעת השמטת אחד המשתנים המסבירים על קריטריון המידע של של Akaike (1974; AIC). במידה והשמטה זו הפחיתה את מדד זה או העלתה אותו בלא יותר מאשר 2 נקודות בהשוואה למודל בעל מדד AIC הנמוך ביותר שהושג עד כה, השמטה זו התקבלה והושמט משתנה מסביר נוסף וחוזר חלילה, עד קבלת המודל הסופי, שכל השמטת משתנה מסביר נוסף ממנו רק תעלה את ערך AIC ב-2 נקודות או יותר.
5. טיב התאמת המודל הסופי לתצפיות נבחנה באמצעות מבחנים וגרפים דיאגנוסטיים הזמינים בחבילה DHARMA (Hartig, 2022) בסביבת R.
6. המובהקות הסטטיסטית (ערך p) של המקדמים של המשתנים המסבירים במודל הסופי נבחנה באמצעות מבחן Wald, כפי שמחושב באמצעות הפונקציה `summary` של המודל.
7. כאשר המודל הסופי כלל מגמה עתית מובהקת, כיוונה (עליה או ירידה) ומידת השינוי (באחוזים ובמספרים מוחלטים – כגון זוחלים לדיגום בחלקה וכו') נקבעו ע"י השוואת תחזיות המודל הסופי לשנת התחלת הניטור ולשנת סיום הניטור ביחידה (כאשר המודל הסופי מצא מגמה עתית מובהקת) עבור דיגומים בעונה ובשעה ממוצעים ליחידה (כאשר העונתיות והדגם היממתי נכללו במודל הסופי). כאשר נמצאה השפעה מובהקת של אחד ממשתני הניטור אך לא נמצאה מגמה עתית מובהקת, תחזיות המודל עבור כל אחת מקטגוריות משתנה הניטור הרלוונטי (כגון קרוב / רחוק מישובים או מחקלאות; כולל השוואת צפון מישור החוף לדרומו, כאשר משתנה זה נכלל במודל הסופי) הושוּוּ זו לזו על מנת לחשב את מידת ההבדל ביניהן (באחוזים ובמספרים מוחלטים). כאשר נמצאה הן מגמה עתית והן השפעה של אחד ממשתני הניטור, הושוּוּ תחזיות הניטור מהשנה הראשונה לשנה האחרונה עבור כל אחת מהקטגוריות של משתנה הניטור הרלוונטי.

- Akaike H (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19 (6): 716–723 <https://doi.org/10.1109/TAC.1974.1100705>
- Brooks ME, Kristensen K, Van Benthem KJ, Magnusson A, Berg CW, Nielsen A, ... & Bolker BM (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal* 9(2): 378–400. <https://doi.org/10.32614/RJ-2017-066>
- Flurry BD, Levri EP (1999). Periodic logistic regression. *Ecology* 80(7): 2254–2260. <https://www.jstor.org/stable/176907>
- Hartig F (2022). DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.6, <https://CRAN.R-project.org/package=DHARMA>
- Lusa L, Ahlin Č (2020). Restricted cubic splines for modelling periodic data. *PLoS ONE* 15(10): e0241364. <https://doi.org/10.1371/journal.pone.0241364>
- R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Urbanek S, Horner J (2022). Cairo: R Graphics Device using Cairo Graphics Library for Creating High-Quality Bitmap (PNG, JPEG, TIFF), Vector (PDF, SVG, PostScript) and Display (X11 and Win32) Output. R package version 1.6–0, <https://CRAN.R-project.org/package=Cairo>
- Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. <https://ggplot2.tidyverse.org/>