Do greener childhoods mean fewer eating disorders? Using satellite imagery to uncover insights

Perla Rivadeneyra[1], Angela Favaro[2], Paolo Meneguzzo[2], Francesco Pirotti[3,4]

[1] Fondazione Eni Enrico Mattei, Venice, Italy - perla.rivadeneyra@feem.it
[2] Department of Neurosciences, University of Padova, Italy - (angela.favaro,paolo.meneguzzo)@unipd.it
[3] Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova, Italy
[4] Department of Land, Environment, Agriculture and Forestry (TESAF), University of Padova, Italy – francesco.pirotti@unipd.it

Keywords: NDVI, Landsat, Health and well-being, Eating disorders, green-spaces

Abstract

The benefits of natural environments on mental health have been documented in numerous studies. In spite of this, the underlying mechanisms by which different elements of natural environments affect mental health, and especially eating disorders, are still largely unknown. In this study, we use remote sensing time series from the Landsat missions to associate a degree of exposure to natural vs urban environment of 659 individuals with diagnosed eating disorders (ED), such as anorexia, bulimia, purging, binge (BED), who constitute our cases. We also obtain data from 39 individuals not affected by eating disorders that comprise our control group. We explore associations between growing up surrounded by areas with a higher NDVI to individuals affected by ED. A Cox regression model was used to estimate rates as relative risk ratios and then adjusted for potential confounding from other known risk factors. The results of the Cox regressions show that the relative risk of developing ED is lower for children growing up in environments with a higher NDVI. Rates did not notably change by adjustment for GDP (a proxy for socioeconomic status), elevation, CO_2 and population. In conclusion, we found that growing up surrounded by greener settings may lower the risk of eating disorders by some 33%. According to our findings, ensuring access to natural environments is crucial for improving public health and creating not only more sustainable but also more livable cities.

1. Introduction

Eating disorders affect 4 to 5% of the Italian population Grave et al. (1997) and contribute to functional impairment, however, few studies have identified risk factors that contribute to the onset of anorexia nervosa (AN), bulimia nervosa (BN), binge eating disorder (BED), and purging disorder (PD). Identifying risk factors specific to each eating disorder is critical for advancing etiologic knowledge and designing effective prevention programs. The epidemiological studies on eating disorders in Italy are available on a few published papers which are Santonastaso et al. (1996) and Ruggiero et al. (2001), with few available studies focusing on Padua including Santonastaso et al. (1996); Favaro et al. (2003). Therefore, the present study is a valuable addition focusing on a specific Italian area in the province of Padua.

There is extensive research exploring various genetic, psychosocial, and cultural factors contributing to the onset of eating disorders Polivy and Herman (2002); Blodgett Salafia et al. (2015); Culbert et al. (2015); Striegel-Moore and Bulik (2007); Cash and Deagle III (1997), and there is a growing recognition of the potential influence of environmental factors, particularly the accessibility and presence of green spaces during the childhood years Hay and Mitchison (2021); Engemann et al. (2019).

Existing evidence has consistently demonstrated the mental health benefits associated with interaction with natural environments, including reduced stress, improved overall well-being, and enhanced psychological resilience Tillmann et al. (2018); Wells (2021); Mantler and Logan (2015). However, the specific impact of childhood exposure to green spaces on the risk of developing eating disorders remains a relatively underexplored terrain.

To investigate this hypothesis, we have adopted the Cox proportional hazards regression model. Cox regression offers an approach to assess the hazard of developing eating disorders over time while accounting for an array of covariates and potential confounding factors Spruance et al. (2004). This methodology allows us to analyze whether individuals who spent their childhood in proximity to green spaces exhibit a reduced risk of developing eating disorders during their later years in the Italian region of Padua, controlling for various demographic, socioeconomic, and one biophysical variable. In this study, we use the mean NDVI (Normalized Difference Vegetation Index) around the residence of each individual to assess the level of green vegetation in the immediate vicinity of their homes during childhood. NDVI is a widely used and well-established measure of vegetation health and density, and it has been employed in various environmental and public health studies Rugel et al. (2017).

In this study we control for an array of covariates, to disentangle the nuanced interactions between childhood green space exposure and the risk of developing eating disorders. In addition to the empirical analysis, we will incorporate pertinent research findings and scientific evidence throughout this study. They will help in contextualizing the significance of our investigation within the broader landscape of environmental determinants of mental health. We use the Cox regression methodology, to provide valuable insights into the potential role of childhood green space exposure in preventing or mitigating the onset of eating disorders, ultimately contributing to the development of targeted interventions and public health initiatives in this domain.

2. Materials and methods

2.1 Data

The first step in the data-gathering process is determining the latitude and longitude of each address for the controls and cases in the database. This is necessary because only addresses are provided, not geolocations. We use the R software to connect to different geocoding services that provide limited free usage and that automatically convert addresses into coordinates. We gauge the amount of perceived green by buffering the NDVI surrounding 1 km of the individual's residence using the coordinates previously obtained in R. We calculate the mean NDVI within the buffer for each available Landsat image the period 1985-2022. The NDVI measures the difference between the amount of light absorbed (red) and reflected (near-infrared) (Engemann et al., 2019), it is calculated in the following way:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

where NIR is the near-infrared and RED is the red band. NDVI is a commonly used and effective measure of green space. Appropriate cloud-masking procedures were carried out to remove as much NDVI values that did not reflect the conditions of the Earth surface in terms of vegetation.

In a study conducted by Reid et al. (2018), the researchers aimed to determine the most effective method for gauging residents' perceptions of greenery in New York. They conducted a comparative analysis of satellite imagery sources, including Landsat, MODIS, and AVHRR, each offering different spatial resolutions, where Landsat offered the highest spatial resolution with a granularity of 30 m. The findings of the study suggested that Landsat imagery closely approximated the way New York residents perceive green spaces. However, the study also emphasized the importance of considering aggregation and buffer size in the assessment. Specifically, it was discovered that for Landsat imagery, a 500-meter buffer around a resident's home yielded the most accurate representation of their perception of greenness. In contrast, for coarser imagery like AVHRR and MODIS, a broader 1-kilometer buffer was found to be more suitable, indicating that residents' perceptions of greenery encompassed a larger area when using lower-resolution imagery. Nevertheless, for the case of Padua we take into account the distinct characteristics of urban spaces with respect to New York, which are likely larger. Therefore, we have chosen to implement a 1-kilometer buffer around residences, recognizing that the buffer size should be tailored to the specific spatial features of the urban environment of Padua.

The NDVI values were further aggregated over time using a time window of 5 years, to further remove noise and smooth the final value to something that represents the situation in terms of how much vegetation is present around each individual. All the procedures were carried out in the Google Earth Engine environment, for efficiency and memory efficiency .

The next step is to assign the NDVI value to each individual considering during the exposure duing her/his childhood. We import the NDVI results into R, where further calculations are performed. All data processing and statistics were performed in R using the packages *readr, dyplr, tidyr, ggplot2, survminer, sf, sp, magittr, lubridate, plyr, raster, rgdal* and *survival*. The first realization is that even if we have individuals who were born before 1985, the Landsat images before that date are very hard to find and to be atmospherically corrected and do not provide reliable results. Therefore, we filter the database for individuals born in and after 1985, with this filtering criteria, the original database with 2390 individuals is reduced to 695, considering both cases and controls. The next step is to create a new aggregated NDVI index during childhood; this way of calculating childhood NDVI was also performed. We did not consider the year 0 of the individual and started the aggregation with the nearest NDVI after this year. All individuals' aggregated NDVI stems from two different years during childhood - where we consider childhood before age 12. The covariates are the mean GDP around 1 km from the surroundings of the individual's residence, the covariate population was also calculated using this same approach, on the other hand, CO_2 was calculated at the smallest administrative level in Italy (comune), the distance to hospital was calculated on QGIS based on direct lines to the hospital (the streets are not considered), year of birth was obtained from the survey and the elevation was calculated precisely at the coordinates of the residence. These calculations were performed in QGIS 3.28.8, and it is important to highlight that all variables were normalized before the regression to render the proportions comparable, in R we use the function *scale()*. Table 1 displays the sources of the variables used in this study.

Variable	Spatial resolution	Source
NDVI	30m	United States Geological Survey
GDP	1km	Wang & Fubao (2023)
CO_2	Municipality	Joint Research Data Centre
Year of Birth	NA	Survey
Distance to hospital	NA	Calculated in QGIS
Elevation	30m	OpenTopography
Population	1km	SEDAC

Table 1. Variable's sources

2.2 Methods

The objective of our study is to investigate whether exposure to green areas (e.g., parks, forests, green spaces) during childhood has any impact on the risk of developing eating disorders in adulthood.

The original idea of this project was to use a spatial econometric method to understand if the spatial component played a role in the association between eating disorders and green exposure during childhood, however, after mapping the cases and controls it soon became evident that the spatial distribution was distorted by a different element than a natural spatial clustering. Thus, by determining the position of the psychiatric hospital of Padua, it became evident that there was a problem with the selection of the sample, since people were choosing to attend the hospital because it was nearby, and only a very few went from more distant places in the Province (see Figure 1). This is a sampling error called auto-selection and does not allow us to properly calculate the relative risk at the municipality level, which excludes using spatial regression or ecological regression. On the other hand, Cox regression has been used in previous studies to examine the association between eating disorders and green exposure during childhood (Engemann et al., 2019, 2020), and for this reason we selected this method for this study as well.

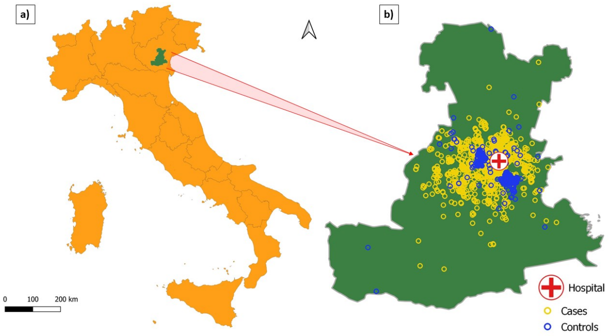

Figure 1. a) Location of the study area: the province of Padua, and nearby municipalities; b) geolocation of cases, controls and the psychiatric Hospital in the area.

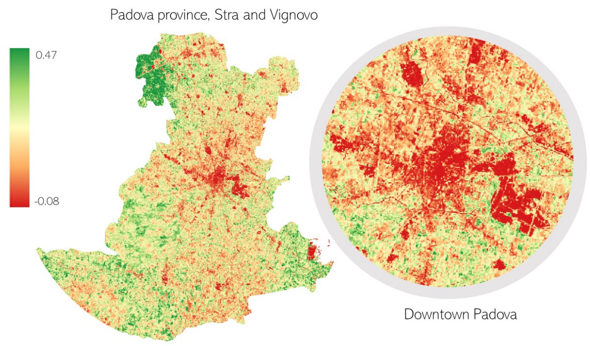

Figure 2. NDVI map of Padua for the year 2022 calculated from Landsat 8.

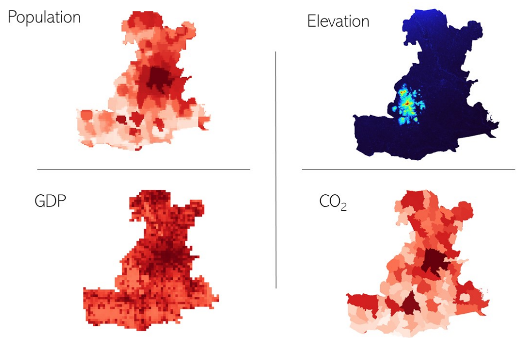

Figure 3. Spatial distribution of the variables used in this study available in raster format.

2.2.1 Cox regression: Cox regression is a statistical method widely employed in fields such as medical research and epidemiology that has a fundamental role in survival analysis. Its purpose is to disentangle the intricate relationship between the survival time of individuals and one or more predictor variables. In essence, Cox regression is a tool for understanding the factors that influence the duration until a specific event of interest takes place. This event can be of any type, ranging from time until the development of symptoms of a disease - which is the way we have calculated the survival time in the present study - or until the failure of a mechanical component or to mortality itself. The main element of the Cox regression model in this study is the concept of the hazard function, it explains how the risk of developing eating disorders changes over time, which is a valuable insight into the dynamics of the development of the disease. By assuming that the hazard rate can be expressed as a function of predictor variables, Cox regression provides a powerful method for exploring and modeling complex survival data, ultimately improving our understanding of critical events and their determinants (Klein et al., 2003; Fox and Weisberg, 2002). The Cox Proportional Hazard Model fits the hazard rate as outlined in the following equation:

$$\lambda(t; x_1, ..., x_p) = \lambda_0(t) \cdot exp(\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + ... + \beta_p \cdot x_p) \quad (2)$$

where $\lambda(t)$ is the hazard rate at time t, $\lambda_0(t)$ is the baseline hazard rate, representing the hazard when all predictor variables are equal to zero, β is the coefficients associated with the predictor variables X_1, X_2, ..., X_p, which measure the log-hazard ratio and exp() the exponential function.

One of the key assumptions of Cox regression is that the hazard ratios for the predictor variables are constant over time, meaning that the effect of the predictors on the hazard rate remains proportional (Hosmer Jr et al., 2011). In other words, it assumes that the hazard (risk of experiencing the event of interest) for one group compared to another group is constant and does not change as time progresses. The proportional hazards assumption is a fundamental assumption in Cox proportional hazards regression, which is commonly used for survival analysis. Mathematically, if we have two groups (e.g. cases vs. control), and we denote the hazard functions for these groups as $h_1(t)$ and $h_2(t)$, respectively, the proportional hazards assumption can be expressed as:

$$\frac{h_1(t)}{h_2(t)} = constant \quad (3)$$

This constant hazard ratio indicates that the relative risk between the two groups is the same at all time points. When the proportional hazards assumption is met, the hazard ratio obtained from the Cox regression analysis is an interpretable measure of relative risk (Kleinbaum and Klein, 1996).

Another element that is relevant for this type of study is being aware of the censoring. In survival analysis, some individuals may not experience the event of interest during the study period. They may develop the disease later on in time. Their survival times are said to be censored because we don’t observe the actual event time but only know that it has not occurred up to a certain point (Hosmer Jr et al., 2011).

Finally, Cox regression models estimate the coefficients using maximum likelihood estimation. These coefficients represent the logarithm of the hazard ratio associated with each predictor variable. *A positive coefficient indicates an increased hazard, while a negative coefficient indicates a decreased hazard.* The estimated coefficients can be used to interpret the effect of the predictor variables on survival time. For example, a hazard ratio greater than 1 suggests an increased risk of the event, while a hazard ratio less than 1 signals a reduced risk (Armitage et al., 2008).

3. Results

In the realm of statistical analysis, Cox proportional hazards regression stands as a robust tool for investigating the factors that influence the timing of specific events. In this context, examining the results obtained from such an analysis sheds light on the significance of various predictor variables and their implications on survival times. Before adding further covariates to the Cox regression we show the model with only NDVI as explanatory variable. In the following two output tables (Tables 3 and 4), we will observe how and to which extent the addition of new variables changes the NDVI coefficient, to test its stability. We use this test as a sensitivity analysis of the NDVI coefficient under varying covariates.

Variable	Coefficient	Hazard Ratio	Std. Error(z)
NDVI	-0.17^{***}	0.84	0.04 (-4.65)
Concordance	0.546		0.01
Likelihood Ratio Test	21.7^{***}		
Wald Test	21.58^{***}		
n	695		

Table 2. Univariate Cox Regression Output: Model 1

In Table 2 the coefficient of NDVI of approximately -0.17 represents the estimated effect of a one-unit change in the scaled NDVI on the hazard of developing an eating disorder (the event). In this context, a negative coefficient indicates that an increase in the scaled NDVI is associated with a decrease in the hazard of the event, it is important to highlight that the univariate model with only NDVI as explanatory variable is statistically significant.

Variable	Coefficient	Hazard Ratio	Standard Error	z
NDVI	-0.39***	0.67	0.06	-6.22
GDP	-0.29***	0.74	0.06	-4.65
CO_2	0.1	1.10	0.11	0.89
birthYear	0.37***	1.44	0.05	7.4
dist_hosp	-0.06	0.94	0.08	-0.8
elevation	0.1*	1.1	0.05	1.9
pop_mean	-0.19	0.82	0.14	-1.34
Concordance	0.63		0.01	
Likelihood Ratio Test	138***			
Wald Test	134***			
Score (logrank)	138***			
n	695			

Table 3. Multivariate Cox Regression Output: Model 2

Each row within Table 3 corresponds to a distinct predictor variable, which are GDP, CO_2, year of birth (birthYear), distance to hospital (dist_hosp), elevation, and population (pop_mean).

The coefficient is the central metric in the table, and it represents the estimated coefficient for each predictor variable, enabling us to quantify the degree of influence that each variable wields.

In column 3, we have included the hazard ratio (HR, or exponentiated coefficient) which constitutes a pivotal component of our analysis. The HR offers valuable insights into how changes in predictor variables can influence the hazard of the event of interest, in our case it is the time to the first appearance of eating disorder symptoms. For instance, the NDVI coefficient is -0.39 and the HR value is 0.59 which signifies that a one-unit increase in the scaled NDVI corresponds to a substantial 33% reduction in the hazard associated with the event - which is the result of calculating: $((1-0.67)\cdot 100)$. The coefficient and hazard ratio of the variable population does not have the expected negative sign, but in this case the coefficient is not significant, and, therefore, we cannot be sure about the true value and true sign of this coefficient.

The standard error (std. error) in column 4, is a metric that provides an understanding of the precision of our coefficient estimates, offering a measure of the inherent uncertainty associated with our findings.

The z-statistic in column 5, is instrumental in quantifying how many standard errors the coefficient estimate deviates from the null hypothesis, aiding in the determination of statistical significance (p-values).

It is important to highlight that our analysis can be extended to the calculation of confidence intervals (CI), specifically the lower .95 and upper .95 values (not available in our output tables). These intervals provide us with a degree of confidence (typically 95%) within which the true hazard ratio is likely to fall. For instance, the 95% confidence interval for NDVI ranges from -0.33 to 0.45, which is calculated by adding the std. error (for the upper CI) and subtracting the std. error for the lower bound.

We also have a measure of the concordance of the survival analysis model. The concordance index summarizes how well a predicted risk score describes an observed sequence of events . The concordance index also known as C-index assesses how well the model distinguishes between individuals who experience the event of interest (e.g., disease onset, mortality) and those who do not. The concordance index (C-index) in Cox regression ranges from 0.5 to 1.0. A value of 0.5 indicates no predictive ability, while 1.0 indicates perfect prediction Kleinbaum and Klein (1996); Harrell et al. (2001). In our table's result, the calculated value of 0.63 indicates a moderate level in the model's ability to correctly rank survival times.

We also have the Likelihood Ratio Test, which compares our model with a null model, it yields a small p-value (p=2e-16), signifying the statistical significance of our model with the included predictors. The Wald Test, assesses the significance of individual predictors within the model, it returns a p-value of p=2e-15, also highlighting the significance of at least one predictor. The Score (Logrank) Test, designed for hypothesis testing in survival analysis, has an extremely small p-value ($p < 2e - 16$) that indicates that the Cox proportional hazards model significantly fits the data. This suggests that the predictors included in the model collectively have a significant impact on the hazard of the event, and the model provides a good fit to the observed survival times.

The results from the Cox regression provide compelling evidence of the significance of our Cox regression model. With NDVI, GDP, CO_2, distance to hospital, elevation, and population as our predictors, our findings underscore the protective effect of exposure to higher NDVI values during childhood, reducing the risk by 33% of developing eating disorders.

https://doi.org/10.5194/isprs-annals-X-3-2024-355-2024 | © Author(s) 2024. CC BY 4.0 License.

Variable	Coef	Hazard Ratio	Std. Error	z
NDVI	-0.49*	0.61	0.07	-6.71
GDP	-0.33**	0.71	0.07	-4.86
CO2	0.08	1.08	0.23	0.34
birthYear	0.36***	1.43	0.05	6.90
dist_hosp	0.15	1.16	0.12	1.20
elevation	0.10*	1.11	0.05	1.94
pop_mean	-0.03	0.97	0.20	-0.16
Municipalities of Padua, and Venetian municipalities of Stra and Vignovo				
Albignasego	-0.15	0.86	0.28	-0.53
Battaglia Terme	0.71	2.04	1.05	0.68
Borgo Veneto	-16.35	0.00	1567.00	-0.01
Cadoneghe	0.02	1.02	0.33	0.06
Campo San Martino	0.91	2.48	1.04	0.88
Casalserugo	-0.38	0.68	0.43	-0.88
Cervarese Santa Croce	-0.07	0.93	0.55	-0.14
Due Carrare	0.84*	2.31	0.43	1.97
Este	-1.02	0.36	1.15	-0.89
Galzignano Terme	-0.06	0.94	0.76	-0.08
Limena	0.90**	2.46	0.41	2.21
Masera Di Padova	-0.55	0.58	0.42	-1.31
Mestrino	-0.68*	0.51	0.37	-1.86
Montegrotto Terme	0.35	1.41	0.36	0.97
Noventa Padovana	-0.22	0.80	0.43	-0.52
Padova	-0.15	0.86	0.39	-0.38
Piombino Dese	-2.49**	0.08	1.08	-2.29
Piove di Sacco	0.46	1.58	1.06	0.43
Ponte di Brenta	0.80	2.22	1.07	0.75
Ponte San Nicola	-0.01	0.99	0.38	-0.03
Rovolon	-0.78	0.46	1.04	-0.75
Rubano	-0.25	0.78	0.30	-0.83
Saccolongo	0.05	1.05	0.57	0.09
Sant Angelo di Piove di Sacco	-0.46	0.63	1.02	-0.44
Saonara	0.10	1.11	0.32	0.31
Selvazzano Dentro	0.12	1.13	0.29	0.42
Solesino	0.20	1.22	1.11	0.18
Stra	-15.41	0.00	1622.00	-0.01
Teolo	0.42*	1.52	0.39	1.07
Torreglia	0.54*	1.71	0.47	1.13
Veggiano	-1.12	0.32	1.03	-1.09
Vigonza	-0.27	0.76	0.74	-0.37
Concordance	0.64		0.01	
Likelihood Ratio Test	189			2×10^{-16}
Wald Test	173			2×10^{-16}
Score (logrank)	192			2×10^{-16}
n	695			
Number of events	656			

Table 4. Multivariate Cox Regression Output: Model 3

In the Cox regression output in Table 4 we have included the different Paduan administrative areas in the regression, it is important to highlight that a few patients live outside the Paduan region, in neighboring areas, these areas are Stra and Vignovo, these areas are considered in maps 1 and 2. The use of the administrative units as a covariate might be relevant to account for municipality-level differences.

In this Cox regression output, we are modeling the survival data with multiple predictors, including both continuous variables and a categorical variable (municipalities of Padua). The results suggest that NDVI, GDP, year of birth, Campo San Martino, Limena, and Piombino Dese have significant negative coefficients. This means that an increase in these variables is associated with a decreased hazard of the event. On the other hand, Due Carrare, elevation, Ponte di Brenta, Solesino, and Torreglia have significant positive coefficients. This implies that an increase in these variables is associated with an increased hazard of the event. The other predictors, such as CO_2, distance to hospital, and many municipalities, do not appear to be statistically significant predictors of the hazard based on their p-values. The Concordance value of 0.643 indicates the model's predictive accuracy, which in this case suggests a moderate predictive performance. The Likelihood Ratio test, Wald test, and Score (logrank) test are all highly significant ($p < 2e-16$), indicating that the overall model is statistically significant.

After examining the three regression outputs - the univariate model, the initial multivariate model, and the extended multivariate model that includes provinces - several trends become evident. First, we notice a progressive decrease in the negative value of the NDVI coefficient as we move from the first to the third model. This suggests that the hazard reduction associated with NDVI increases as we introduce additional variables into the model.

Furthermore, it is noteworthy that in all three models, the NDVI coefficient remains statistically significant, although its significance level diminishes in the third model. This indicates that NDVI continues to have a notable impact on the hazard, but other factors introduced in the third model might be influencing its significance. After considering these observations, we lean towards the second model as the most reliable. The third model, due to its inclusion of numerous additional coefficients related to municipalities, could potentially introduce multicollinearity issues or lead to inflated coefficients, including the NDVI coefficient. Therefore, the second model is more balanced in terms of complexity and reliability in capturing the relationship between NDVI and developing eating diroders.

4. Discussion

The analyzed database has enormous potential for improved and more in-depth studies of the etiology of eating disorders. First, we would suggest calculating a more reliable NDVI for satellites before Landsat 5 (or NDVI's before 1985), which is the last Landsat satellite that offers top-of-the-atmosphere (TOA) corrected images. Due to time constraints, we did not atmospherically correct Landsat images before Landsat 5, if done, this would allow the analysis of a big portion of controls in the database. Even though we calculated the NDVI beginning from 1972 (Landsat 1), the NDVI results cannot be relied on, since the images lack the already mentioned TOA correction.

Additionally, it is important to note that while the survey offers valuable insights, there are opportunities for enhancement by considering additional variables. For example, incorporating information about family history in relation to the disease can provide a more comprehensive understanding of the interplay between genetics and environmental factors in the development of eating disorders. Also, incorporating data on parental education and the educational background of the respondents themselves can offer valuable insights into the role of education in overall well-being and its potential impact on eating disorder development or prevention. Additionally, gathering information on the individual's mental health history can provide us with a deeper understanding of how past mental health experiences may influence the onset and progression of eating disorders.

Furthermore, considering gender as a covariate would be an interesting dimension to take into account. Research by Astell-Burt et al. (2014) showed that the connection between green spaces and mental health can differ between men and women. They discovered that women, at certain ages, are more likely to be influenced positively by green spaces compared to men.

Expanding our investigation to encompass a variety of natural settings, including blue landscapes formed by water bodies, presents an exciting avenue for research. By examining the influence of these diverse natural environments on mental health, can give a more holistic understanding of how childhood environments may shape eating disorders in adulthood. This can also be seen as an ecosystem service that forests bring to nearby communities and that must be protected, especially if vulnerable to the effects of climate change Kanan et al. (2023).

Another aspect worth noting is that phenology of green areas has a specific seasonality that might be also be a necessary part of the effect that it brings to the population. Remote sensing can monitor also the phenology thanks to higher availability of open data sources of imagery E. Kosczor et al. (2022). A close relationship to the state of the green canopies is also related to pathogens that might change the availability of healthy plants Dalponte et al. (2023); this aspect is not considered, but might be significant.

This broader perspective allows us to explore the positive impact of natural settings on mental well-being and potentially uncover strategies for enhancing mental health through environmental design.

In summary, while the current survey provides valuable insights, there is great potential for enriching our understanding by incorporating additional variables and by calculating the NDVI for years before 1985. Exploring family history, education, mental health, and gender as key factors can lead to more comprehensive insights. Furthermore, broadening our study to encompass various natural settings opens doors to understanding the potential long-term benefits of childhood environments on adult mental health outcomes.

5. Conclusions

The obtained Cox regression models consistently underscore the significance of NDVI (Normalized Difference Vegetation Index) in relation to the hazard of developing eating disorders. This analysis provides valuable insights into the role of natural environments in mental health, in particular eating disorders.

In Table 2, where we initially examine the univariate model with only NDVI as the explanatory variable, the negative NDVI coefficient of approximately -0.17 signifies a critical finding. It suggests that an increase in the scaled NDVI, representing higher green space exposure, is associated with a decrease in the hazard of developing eating disorders. In essence, this implies that individuals who had greater access to green spaces during childhood may have a lower risk of developing eating disorders in later life. Importantly, this finding is statistically significant, emphasizing the importance of natural environments in promoting mental health. However, a univariate model offers lower levels of reliability, a multivariate context is more robust, therefore we use this first model only as a benchmark model.

As we progress to Table 3, introducing multiple predictor variables into the model, NDVI's significance remains intact. The NDVI coefficient, of -0.39, reinforces the protective effect of green spaces. The hazard ratio of 0.67 translates to a substantial 33% reduction in the hazard for every one-unit increase in scaled NDVI. This finding reaffirms that exposure to green spaces not only remains a significant factor but also retains its protective influence, even when considering other potential variables that may affect the development of eating disorders.

Furthermore, in Table 4, where we incorporate different Paduan administrative areas into the model, we continue to observe the stability and significance of the NDVI coefficient. Despite the added complexity of considering specific geographical locations and potential variations in green spaces across these areas, the NDVI coefficient remains negative at -0.49 and statistically significant. This reiteration of the protective effect of green spaces highlights its robustness as a contributing factor to mental well-being. The hazard ratio of 0.61 indicates a 39% reduction in the hazard for each one-unit increase in scaled NDVI, further emphasizing its role in reducing the risk of eating disorders.

In conclusion, our Cox regression analysis consistently demonstrates that exposure to green spaces, as measured by NDVI, plays a pivotal role in reducing the hazard of developing eating disorders. This finding suggests that promoting access to natural environments during childhood may contribute significantly to mental health and well-being in adulthood. Even as we account for various potential confounding factors and geographical differences, the protective effect of green spaces remains a strong and statistically significant predictor.

The implications of these results advocate for policies and interventions aimed at increasing access to green spaces, particularly during the formative years of childhood. Such efforts may have a positive impact on mental health outcomes and potentially reduce the incidence of eating disorders. The second model, with its inclusion of multiple predictors and the consistent significance of NDVI, emerges as a reliable framework for further exploration of the relationship between natural environments and mental well-being.

References

Armitage, P., Berry, G., Matthews, J. N. S., 2008. *Statistical methods in medical research*. John Wiley & Sons.

Astell-Burt, T., Mitchell, R., Hartig, T., 2014. The association between green space and mental health varies across the lifecourse. A longitudinal study. *J Epidemiol Community Health*, 68(6), 578–583.

Blodgett Salafia, E. H., Jones, M. E., Haugen, E. C., Schaefer, M. K., 2015. Perceptions of the causes of eating disorders: a comparison of individuals with and without eating disorders. *Journal of eating disorders*, 3(1), 1–10.

Cash, T. F., Deagle III, E. A., 1997. The nature and extent of body-image disturbances in anorexia nervosa and bulimia nervosa: A meta-analysis. *International journal of eating disorders*, 22(2), 107–126.

Culbert, K. M., Racine, S. E., Klump, K. L., 2015. Research Review: What we have learned about the causes of eating disorders–a synthesis of sociocultural, psychological, and biological research. *Journal of Child Psychology and Psychiatry*, 56(11), 1141–1164.

Dalponte, M., Cetto, R., Marinelli, D., Andreatta, D., Salvadori, C., Pirotti, F., Frizzera, L., Gianelle, D., 2023. Spectral separability of bark beetle infestation stages: A single-tree time-series analysis using Planet imagery. *Ecological Indicators*, 153.

E. Kosczor, Forkel, M., Hernández, J., Kinalczyk, D., Pirotti, F., Kutchartt, E., 2022. Assessing land surface phenology in Araucaria-Nothofagus forests in Chile with Landsat 8/Sentinel-2 time series. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102862. https://linkinghub.elsevier.com/retrieve/pii/S1569843222000644.

Engemann, K., Pedersen, C. B., Arge, L., Tsirogiannis, C., Mortensen, P. B., Svenning, J.-C., 2019. Residential green space in childhood is associated with lower risk of psychiatric disorders from adolescence into adulthood. *Proceedings of the national academy of sciences*, 116(11), 5188–5193.

Engemann, K., Svenning, J.-C., Arge, L., Brandt, J., Erikstrup, C., Geels, C., Hertel, O., Mortensen, P. B., Plana-Ripoll, O., Tsirogiannis, C. et al., 2020. Associations between growing up in natural environments and subsequent psychiatric disorders in Denmark. *Environmental research*, 188, 109788.

Favaro, A., Ferrara, S., Santonastaso, P., 2003. The spectrum of eating disorders in young women: a prevalence study in a general population sample. *Psychosomatic Medicine*, 65(4), 701–708.

Fox, J., Weisberg, S., 2002. Cox proportional-hazards regression for survival data. *An R and S-PLUS companion to applied regression*, 2002.

Grave, R. D., De Luca, L., Oliosi, M., 1997. Eating attitudes and prevalence of eating disorders: A survey in secondary schools in Lecce, southern Italy. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 2, 34–37.

Harrell, F. E. et al., 2001. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. 608, Springer.

Hay, P., Mitchison, D., 2021. Urbanization and eating disorders: a scoping review of studies from 2019 to 2020. *Current opinion in psychiatry*, 34(3), 287–292.

Hosmer Jr, D. W., Lemeshow, S., May, S., 2011. *Applied survival analysis: regression modeling of time-to-event data*. 618, John Wiley & Sons.

Kanan, A. H., Pirotti, F., Masiero, M., Rahman, M. M., 2023. Mapping inundation from sea level rise and its interaction with land cover in the Sundarbans mangrove forest. *Climatic Change*, 176(8), 104. https://link.springer.com/10.1007/s10584-023-03574-5.

Klein, J. P., Moeschberger, M. L. et al., 2003. *Survival analysis: techniques for censored and truncated data*. 1230, Springer.

Kleinbaum, D. G., Klein, M., 1996. *Survival analysis a self-learning text*. Springer.

Mantler, A., Logan, A. C., 2015. Natural environments and mental health. *Advances in Integrative Medicine*, 2(1), 5–12.

Polivy, J., Herman, C. P., 2002. Causes of eating disorders. *Annual review of psychology*, 53(1), 187–213.

Reid, C. E., Kubzansky, L. D., Li, J., Shmool, J. L., Clougherty, J. E., 2018. It’s not easy assessing greenness: a comparison of NDVI datasets and neighborhood types and their associations with self-rated health in New York City. *Health & place*, 54, 92–101.

Rugel, E. J., Henderson, S. B., Carpiano, R. M., Brauer, M., 2017. Beyond the Normalized Difference Vegetation Index (NDVI): developing a natural space index for population-level health research. *Environmental research*, 159, 474–483.

Ruggiero, G. M., Prandin, M., Mantero, M., 2001. Eating disorders in Italy: A historical review. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association*, 9(5), 292–300.

Santonastaso, P., Zanetti, T., Sala, A., Favaretto, G., Vidotto, G., Favaro, A., 1996. Prevalence of eating disorders in Italy: a survey on a sample of 16-year-old female students. *Psychotherapy and Psychosomatics*, 65(3), 158–162.

Spruance, S. L., Reid, J. E., Grace, M., Samore, M., 2004. Hazard ratio in clinical trials. *Antimicrobial agents and chemotherapy*, 48(8), 2787–2792.

Striegel-Moore, R. H., Bulik, C. M., 2007. Risk factors for eating disorders. *American psychologist*, 62(3), 181.

Tillmann, S., Tobin, D., Avison, W., Gilliland, J., 2018. Mental health benefits of interactions with nature in children and teenagers: A systematic review. *J Epidemiol Community Health*, 72(10), 958–966.

Wells, N. M., 2021. The natural environment as a resilience factor: Nature’s role as a buffer of the effects of risk and adversity. *Nature and psychology: Biological, cognitive, developmental, and social pathways to well-being*, 195–233.