DENVar: A global Dynamic Estimate of NDVI Variance

Stefano Mezzini[1,2] Michael J. Noonan[1,2,3,*]

[1] Okanagan Institute for Biodiversity, Resilience, and Ecosystem Services, The University of British Columbia Okanagan, Kelowna, British Columbia, Canada.

[2] Department of Biology, The University of British Columbia Okanagan, Kelowna, British Columbia, Canada.

[3] Department of Computer Science, Math, Physics, and Statistics, The University of British Columbia Okanagan, Kelowna, British Columbia, Canada.

[*] Correspondence: Michael J. Noonan <michael.noonan@ubc.ca>

Article type: Research article

Words in abstract: 262

Words in main text: 1388

Figures: XXX

Tables: XXX

References: XXX

Appendices: XXX

Key words: remote sensing, NDVI, environmental stochasticity, XXX

Abstract

Since the 1980s, remote-sensing estimates of habitat productivity such as the Normalized Differnece Vegetation Index (NDVI) have helped develop and support a variety of hypotheses and inform policy and industriy-related decisions. The free availability of such data has allowed widespread access and use of well-established metrics of habitat productivity. However, historically, users tended to focus solely on (predictable) trends in productivity. In recent years, people have been recognizing the importance of envrironmental unpredictability (stochasticity), particularly as climate change and human-induced rapid environmental change transform the ecosystems species have evolved in and adapted to. We adress this gap by presenting a new global Dynamic Estimate of NDVI Variance, *DENVar*. We estimate spatiotemporal trends in mean NDVI and the variance around the mean using Hierarchical Generalized Additive Models, which provide a flexible yet transparent model structure, unlike machine learning approaches such as neural netwoks. We show that DENVar can be used as a reliable proxy of environmental stochasticity, strongly correlates with ***XXX***, and can be used to test many hepotheses related to environmental stochasitcity, such as forage unpredictability, regime shifts, ... We conclude by offering some considerations around quantifying and interpreting envirnomental stochasticity, and we provide a publicly and freely available shiny app for estimating mean NDVI and the variance around it, DENVar, for given coordinates and optional dates. The app requires no knowledge of statistics or coding. Additionally, we provide a temporally static raster of mean NDVI and DENVar for the years 1981-2025 that can be used for GIS applications and does not require any knowledge of `R`.

Introduction

Developments in remote sensing since the 1970s have greatly expanded our ability to quantify and monitor landscapes using satellite imagery

- NDVI is an estimate of greenness, NPP, forage availability
- NDVI has many uses and is often convenient because it's free, well-established, and

many people focus on changes in the mean

ability to predict and respond to changes in mean conditions (forage availability, phenology, etc.) depends on the scale of changes relative to the perception scale (Levin, 1992; Frankham & Brook, 2004; Riotte-Lambert & Matthiopoulos, 2020; Steixner-Kumar & Gläscher, 2020) and memory (Abrahms *et al.*, 2019). Consequently, environmental stochasticity is also an important driver, besides changes in the mean (Mezzini *et al.*, 2025)

- variance is hard to calculate, as it depends on a correct estimate of the mean
- paper on estimating the asymptote of environmental variance globally, but variance for a closed system must converge towards a finite value (Bachmaier & Backes, 2011; Fleming *et al.*, 2014)

- we present DENVar
- HGAMs are higly flexible models that allow us to estimate trends in the data without imposing preconceived expectations
- more transparent than black-box machine learning methods like neural networks, etc.
- https://www.sciencedirect.com/science/article/abs/pii/S0034425718305625
- https://silvis.forest.wisc.edu/data/dhis/
- we provide global rasters of the mean and variance in NDVI, averaged across 1981-2025
- we provide a shiny app for people to calculate mean and variance without any knowledge of `R`

Methods

Choice of color schemes

We represent NDVI using a modified version of Fabio Crameri's divergent *bukavu* palette (Crameri, 2018), which has high-contrast for deuteranope and protanope vision. For NDVI values between 0 and 1 (the right half of the palette), the colors also have sufficient contrast for the colors to be distinguishable by both tritanope and achomatic vision. Appendix A contains an approximate representation of the color palette for each vision type. We obtained all color palettes from the `khroma` package (v. 1.14.0, Frerebeau, 2024) for `R` (v. 4.4.1, R Core Team, 2024).

Input data

We obtained NDVI data using the image composites from the AVHRR and VIIRS sensors (Vermote & NOAA CDR Program, 2018, 2022). The code for downloading the data directly from the NOAA server is available at https://github.com/QuantitativeEcologyLab/ndvi-stochasticity/blob/main/analysis/002-download-ndvi-rasters.R. The number of non-NA land raster cells was substantially larger than the maximum data frame size in `R` ($2.28 \times 10^{10} > 2^{31} - 1 \approx 0.21 \times 10^{10}$), so we reduced the dataset size by calculating 15-day averages of NDVI (see Fig. A2) and aggregating the averaged rasters by a factor of 2×2 with `terra::aggregate(fact = 2)`. Although this resulted in a reduction in sample size, the temporal averaging also reduced signal-related noise.

The predictor data for the models included: WWF biome (Olson & Dinerstein, 2002, see Fig. A3), ecoregion (the individual polygons from Olson & Dinerstein, 2002, see Fig. A4), integer day of year (1 to 366), integer year (1981 to 2025), and elevation above sea level (Fig. A5). We excluded all polygons that did not have any NDVI data (0.06% of land, 29% of the polygons, maximum area: 452 km$^2 \approx (21\text{ km})^2$). We distinguished between biomes in the northern and southern hemisphere to allow for different seasonal trends between hemispheres.

The 15-day averages allowed us to have non-repeating values of day of year across years, since 15 is not a factor of 365 or 366. We downloaded the global digital elevation model using the `get_elev_raster()` function from the `elevatr` package for `R` (v. 0.99.0, Hollister *et al.*, 2023) with a resolution of 0.076 degrees.

Modeling

We estimated spatiotemporal trends in mean NDVI using Hierarchical Generalized Additive Models (HGAMs) via the `mgcv` package for `R` (v. 1.9-3, Wood, 2017). To reduce modeling fitting times with negligible losses to model accuracy, we used the `bam()` function with fast REstricted Marginal Likelihood (`method = fREML`) and covariate discretization (`discrete = TRUE`). See Wood, Goude & Shaw (2015) and Wood *et al.* (2017) for more information. Ideally, NDVI should be modeled using beta location-scale models (after the linear transformation $Y^* = \frac{Y+1}{2}$) to account for the fact that: (1) NDVI is bounded between -1 and 1, and (2) the variance in NDVI is dependent on the mean (and vice-versa), since ecosystems with very low NDVI (e.g., rock, ice, or concrete) or very high NDVI (e.g., dense forest) tend to have lower variance in NDVI. However, preliminary tests showed that fitting times for beta models were prohibitive, especially for beta location-scale models. Additionally, Gaussian models fit substantially faster and provided very similar spatialtemporal estimates of mean NDVI (see the `analysis/000-sardinia-test.R` script in the GitHub repository). The model for the mean NDVI had the structure below:

```
m_mean <- bam(
  ndvi_15_day_mean ~
    biome + # to avoid intercept shrinkage
    s(poly_id, bs = 'mrf', xt = list(nb = nbs)) +
    s(doy, biome, bs = 'fs', xt = list(bs = 'cc'), k = 10) +
    s(year, biome, bs = 'fs', xt = list(bs = 'cr'), k = 10) +
    ti(doy, year, biome, bs = c('cc', 'cr', 're'), k = c(5, 5)) +
    s(elevation_m, bs = 'cr', k = 5),
  family = gaussian(),
  data = d,
  method = 'fREML',
  knots = list(doy = c(0.5, 366.5)),
  drop.unused.levels = TRUE,
  discrete = TRUE,
  samfrac = 0.001, # find initial guesses with a subset of the data
  nthreads = future::availableCores(logical = FALSE) - 2,
  control = gam.control(trace = TRUE))
```

Biome-specific fixed-effect intercepts allowed to account for the differences in mean NDVI throught biomes without coefficient shrinkage (i.e., reversion towards the global mean). A Markov Random Field of the ecoregions (`bs = mrf`, see page ***240?*** of Wood, 2017) estimated smoothed, ecoregion-specific deviations from the biome-level spatial means. Two factor-smooth interaction terms (`bs = 'fs'`) accounted for biome-specific seasonal and yearly trends. Seasonal terms were made cyclical by using cyclical cubic splines (`xt = list(bs = 'cc')`), which are continuous up to and including second derivative at the edge knots (0.5 and 366.5, which correspond to 00:00 of January 1^{st} and 24:00 of December 31^{st}, respectively). A tensor product interaction smooth for each biome estimated the change in seasonal trends over the years. Finally, a smooth of elevation above sea level accounted for the effect of altitude. See Pedersen *et al.* (2019) for more information on hierarchical modeling and the use and interpretation of factor smooth interaction terms.

will we subtract the average residual from each pixel?

A second HGAM estimated the variance in NDVI around the mean estimated by the model above. The model had an identical structure to the model for the mean, with the exception that (1) the response variable was the squared residual from the first model, such that the model estimated the mean squared residual (i.e., the variance) for a given point in time and space, and (2) a smooth term of the estimated mean NDVI. The model is available below:

```r
m_var <- bam(
  e_2 ~
    biome +
    s(poly_id, bs = 'mrf', xt = list(nb = nbs)) +
    s(doy, biome, bs = 'fs', xt = list(bs = 'cc'), k = 10) +
    s(year, biome, bs = 'fs', xt = list(bs = 'cr'), k = 10) +
    ti(doy, year, biome, bs = c('cc', 'cr', 're'), k = c(5, 5)) +
    s(elevation_m, bs = 'cr', k = 5) +
    s(mu_hat, bs = 'cr', k = 5),
  family = gaussian(),
  data = d,
  method = 'fREML',
  knots = list(doy = c(0.5, 366.5)),
  drop.unused.levels = TRUE,
  discrete = TRUE,
  samfrac = 0.001,
```

```r
  nthreads = future::availableCores(logical = FALSE) - 2,
  control = gam.control(trace = TRUE))
```

Results

output of `draw()`

hex plots of mean and var vs common metrics:

- temperature
- seasonal temperature range
- precip
- seasonal temperature range
- spp diversity
- gross primary productivity
- hfi
- max animal weight
- https://silvis.forest.wisc.edu/globalwui/
- https://silvis.forest.wisc.edu/data/dhis/: NDVI16 (Normalized Difference Vegetation Index), EVI16 (Enhanced Vegetation Index), FPAR8 (Fraction absorbed Photosynthetically Active Radiation), LAI8 (Leaf Area Index), GPP8 (Gross Primary Productivity)
- https://silvis.forest.wisc.edu/data/eurasia-trends/

Discussion

- green wave surfing: change in the mean, failure to surf: variance
- migration routes are key habitat (Ortega et al., in prep), especially in stochastic environments
- Sawyer *et al.* (2019)

- https://www.sciencedirect.com/science/article/pii/S0960982220308484
- https://onlinelibrary.wiley.com/doi/full/10.1111/gcb.15169
- industrial development affects animals' ability to respond to both predictable and unpredictable change (https://www.nature.com/articles/s41559-022-01887-9)
- effects of sudden deep snow in red desert on proghorn (https://www.sciencedirect.com/science/article/abs/pii/S0960982225002957)
- https://www.nature.com/articles/s41467-023-37750-z
- https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/ecy.4238
- see data for snow and pronghorn paper (Aikens et al. 2025)
- detecting regime shifts (in lakes: Bjorndahl *et al.*, 2022)
- smoothness of mu impacts smoothness and size of s2: depends on animals' ability to respond to, learn, and predict E(R)

Why not use machine learning? / Advantages of not using ML

arguably, GAMs can be considered as a form of machine learning, but we are referring to neural networks and other forms of black-box ML.

Useful references

- `{khroma}`: Frerebeau (2024)
- Keith *et al.* (2008)
- Chevin, Lande & Mace (2010)
- Rickbeil *et al.* (2019)
- Mueller *et al.* (2008)
- Pettorelli *et al.* (2005)
- Keys, Barnes & Carter (2021)
- Nilsen, Herfindal & Linnell (2005)
- Stephens & Charnov (1982)
- Merkle *et al.* (2016)
- Tian *et al.* (2015)
- Huang *et al.* (2021)
- Fan & Liu (2016)
- Wang *et al.* (2019)
- Pease (2024)
- Xu *et al.* (2021)
- Mezzini *et al.* (2025)
- **TO READ:** Site fidelity as a maladaptive behavior in the Anthropocene: https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/fee.2456

References

Abrahms B., Hazen E.L., Aikens E.O., Savoca M.S., Goldbogen J.A., Bograd S.J., *et al.* (2019). Memory and resource tracking drive blue whale migrations. *Proceedings of the National Academy of Sciences* **116**, 5582–5587. https://doi.org/10.1073/pnas.1819031116

Bachmaier M. & Backes M. (2011). Variogram or semivariogram? Variance or semivariance? Allan variance or introducing a new term? *Mathematical Geosciences* **43**, 735–740. https://doi.org/10.1007/s11004-011-9348-3

Bjorndahl J.A., Gushulak C.A.C., Mezzini S., Simpson G.L., Haig H.A., Leavitt P.R., *et al.* (2022). Abrupt changes in the physical and biological structure of endorheic upland lakes due to 8-m lake-level variation during the 20 $^{\text{th}}$ century. *Limnology and Oceanography* **67**, 1022–1039. https://doi.org/10.1002/lno.12054

Chevin L.-M., Lande R. & Mace G.M. (2010). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *PLoS Biology* **8**, e1000357. https://doi.org/10.1371/journal.pbio.1000357

Crameri F. (2018). Geodynamic diagnostics, scientific visualisation and StagLab 3.0. *Geoscientific Model Development* **11**, 2541–2562. https://doi.org/10.5194/gmd-11-2541-2018

Fan X. & Liu Y. (2016). A global study of NDVI difference among moderate-resolution satellite sensors. *ISPRS Journal of Photogrammetry and Remote Sensing* **121**, 177–191. https://doi.org/10.1016/j.isprsjprs.2016.09.008

Fleming C.H., Calabrese J.M., Mueller T., Olson K.A., Leimgruber P. & Fagan W.F. (2014). From fine-scale foraging to home ranges: A semivariance approach to identifying movement modes across spatiotemporal scales. *The American Naturalist* **183**, E154–E167. https://doi.org/10.1086/675504

Frankham R. & Brook B.W. (2004). The importance of time scale in conservation biology and ecology. *Annales Zoologici Fennici* **41**, 459–463

Frerebeau N. (2024). *Khroma: Colour schemes for scientific data visualization*. Université Bordeaux Montaigne, Pessac, France.

Hollister J., Shah T., Nowosad J., Robitaille A.L., Beck M.W. & Johnson M. (2023). *Elevatr: Access elevation data from various APIs*.

Huang S., Tang L., Hupy J.P., Wang Y. & Shao G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research* **32**, 1–6. https://doi.org/10.1007/s11676-020-01155-1

Keith D.A., Akçakaya H.R., Thuiller W., Midgley G.F., Pearson R.G., Phillips S.J., *et al.* (2008). Predicting extinction risks under climate change: Coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters* **4**, 560–563. https://doi.org/10.1098/rsbl.2008.0049

Keys P.W., Barnes E.A. & Carter N.H. (2021). A machine-learning approach to human footprint index estimation with applications to sustainable development. *Environmental Research Letters* **16**, 044061. https://doi.org/10.1088/1748-9326/abe00a

Levin S.A. (1992). The problem of pattern and scale in ecology: The robert h. MacArthur award lecture. *Ecology* **73**, 1943–1967. https://doi.org/10.2307/1941447

Merkle J.A., Monteith K.L., Aikens E.O., Hayes M.M., Hersey K.R., Middleton A.D., *et al.* (2016). Large herbivores surf waves of green-up during spring. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20160456. https://doi.org/10.1098/rspb.2016.0456

Mezzini S., Fleming C.H., Medici E.P. & Noonan M.J. (2025). How resource abundance and resource stochasticity affect organisms' range sizes. *Movement Ecology* **13**, 20. https://doi.org/10.1186/s40462-025-00546-5

Mueller T., Olson K.A., Fuller T.K., Schaller G.B., Murray M.G. & Leimgruber P. (2008). In search of forage: Predicting dynamic habitats of mongolian gazelles using satellite-based estimates of vegetation productivity. *Journal of Applied Ecology* **45**, 649–658. https://doi.org/10.1111/j.1365-2664.2007.01371.x

Nilsen E.B., Herfindal I. & Linnell J.D.C. (2005). Can intra-specific variation in carnivore home-range size be explained using remote-sensing estimates of environmental productivity? *Écoscience* **12**, 68–75. https://doi.org/10.2980/i1195-6860-12-1-68.1

Olson D.M. & Dinerstein E. (2002). The global 200: Priority ecoregions for global conservation. *Annals of the Missouri Botanical Garden* **89**, 199. https://doi.org/10.2307/3298564

Pease B.S. (2024). Ecological scales of effect vary across space and time. *Ecography* **2024**, e07163. https://doi.org/10.1111/ecog.07163

Pedersen E.J., Miller D.L., Simpson G.L. & Ross N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ* **7**, e6876. https://doi.org/10.7717/peerj.6876

Pettorelli N., Vik J.O., Mysterud A., Gaillard J.-M., Tucker C.J. & Stenseth N.Chr. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution* **20**, 503–510. https://doi.org/10.1016/j.tree.2005.05.011

R Core Team (2024). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rickbeil G.J.M., Merkle J.A., Anderson G., Atwood M.P., Beckmann J.P., Cole E.K., *et al.* (2019). Plasticity in elk migration timing is a response to changing environmental conditions. *Global Change Biology* **25**, 2368–2381. https://doi.org/10.1111/gcb.14629

Riotte-Lambert L. & Matthiopoulos J. (2020). Environmental predictability as a cause and consequence of animal movement. *Trends in Ecology & Evolution* **35**, 163–174. https://doi.org/10.1016/j.tree.2019.09.009

Sawyer H., Merkle J.A., Middleton A.D., Dwinnell S.P.H. & Monteith K.L. (2019). Migratory plasticity is not ubiquitous among large herbivores. *Journal of Animal Ecology* **88**, 450–460. https://doi.org/10.1111/1365-2656.12926

Steixner-Kumar S. & Gläscher J. (2020). Strategies for navigating a dynamic world. *Science* **369**, 1056–1057. https://doi.org/10.1126/science.abd7258

Stephens D.W. & Charnov E.L. (1982). Optimal foraging: Some simple stochastic models. *Behavioral Ecology and Sociobiology* **10**, 251–263. https://doi.org/10.1007/BF00302814

Tian F., Fensholt R., Verbesselt J., Grogan K., Horion S. & Wang Y. (2015). Evaluating temporal consistency of long-term global NDVI datasets for trend analysis. *Remote Sensing of Environment* **163**, 326–340. https://doi.org/10.1016/j.rse.2015.03.031

Vermote E. & NOAA CDR Program (2018). NOAA climate data record (CDR) of AVHRR normalized difference vegetation index (NDVI), version 5

Vermote E. & NOAA CDR Program (2022). NOAA climate data record (CDR) of VIIRS normalized difference vegetation index (NDVI), version 1

Wang X., Cao L., Fox A.D., Fuller R., Griffin L., Mitchell C., *et al.* (2019). Stochastic simulations reveal few green wave surfing populations among spring migrating herbivorous waterfowl. *Nature Communications* **10**, 2187. https://doi.org/10.1038/s41467-019-09971-8

Wood S.N. (2017). *Generalized additive models: An introduction with r*, Second edition. CRC Press/Taylor & Francis Group, Boca Raton.

Wood S.N., Goude Y. & Shaw S. (2015). Generalized additive models for large data sets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **64**, 139–155. https://doi.org/10.1111/rssc.12068

Wood S.N., Li Z., Shaddick G. & Augustin N.H. (2017). Generalized additive models for gigadata: Modeling the u.k. Black smoke network daily data. *Journal of the American Statistical Association* **112**, 1199–1210. https://doi.org/10.1080/01621459.2016.1195744

Xu W., Barker K., Shawler A., Van Scoyoc A., Smith J.A., Mueller T., *et al.* (2021). The plasticity of ungulate migration in a changing world. *Ecology* **102**, e03293. https://doi.org/10.1002/ecy.3293