

RESEARCH ARTICLE

# A Changing Number of Alternative States in the Boreal Biome: Reproducibility Risks of Replacing Remote Sensing Products

Chi Xu<sup>1,2\*</sup>, Milena Holmgren<sup>3</sup>, Egbert H. Van Nes<sup>2</sup>, Marina Hirota<sup>4</sup>, F. Stuart Chapin, III<sup>5</sup>, Marten Scheffer<sup>2</sup>

**1** School of Life Sciences, Nanjing University, Xianlin Road 163, Nanjing, 210023, P.R. China, **2** Aquatic Ecology and Water Quality Management Group, Wageningen University, P.O. Box 47, NL-6700 AA, Wageningen, The Netherlands, **3** Resource Ecology Group, Wageningen University, P.O. Box 47, NL-6700 AA, Wageningen, The Netherlands, **4** Department of Physics, Federal University of Santa Catarina, P.O. Box 476, 88040–970, Florianópolis, Brazil, **5** Institute of Arctic Biology, University of Alaska, Fairbanks, Alaska, 99775, United States of America

\* [xuchi@nju.edu.cn](mailto:xuchi@nju.edu.cn)



OPEN ACCESS

**Citation:** Xu C, Holmgren M, Van Nes EH, Hirota M, Chapin FS, III, Scheffer M (2015) A Changing Number of Alternative States in the Boreal Biome: Reproducibility Risks of Replacing Remote Sensing Products. PLoS ONE 10(11): e0143014. doi:10.1371/journal.pone.0143014

**Editor:** Shijo Joseph, Center for International Forestry Research (CIFOR), INDONESIA

**Received:** May 11, 2015

**Accepted:** October 29, 2015

**Published:** November 16, 2015

**Copyright:** © 2015 Xu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The remotely-sensed data of tree cover are available from the Global Land Cover Facility website (<http://www.landcover.org/data/>).

**Funding:** This work was partly funded by the European Research Council (ERC) Advanced Grant and Spinoza award to MS. CX was supported by National Natural Science Foundation of China (41271197) and the China Scholarship Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Publicly available remote sensing products have boosted science in many ways. The openness of these data sources suggests high reproducibility. However, as we show here, results may be specific to versions of the data products that can become unavailable as new versions are posted. We focus on remotely-sensed tree cover. Recent studies have used this public resource to detect multi-modality in tree cover in the tropical and boreal biomes. Such patterns suggest alternative stable states separated by critical tipping points. This has important implications for the potential response of these ecosystems to global climate change. For the boreal region, four distinct ecosystem states (i.e., treeless, sparse and dense woodland, and boreal forest) were previously identified by using the Collection 3 data of MODIS Vegetation Continuous Fields (VCF). Since then, the MODIS VCF product has been updated to Collection 5; and a Landsat VCF product of global tree cover at a fine spatial resolution of 30 meters has been developed. Here we compare these different remote-sensing products of tree cover to show that identification of alternative stable states in the boreal biome partly depends on the data source used. The updated MODIS data and the newer Landsat data consistently demonstrate three distinct modes around similar tree-cover values. Our analysis suggests that the boreal region has three modes: one sparsely vegetated state (treeless), one distinct ‘savanna-like’ state and one forest state, which could be alternative stable states. Our analysis illustrates that qualitative outcomes of studies may change fundamentally as new versions of remote sensing products are used. Scientific reproducibility thus requires that old versions remain publicly available.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

A recent call for more transparent, open and reproducible science stressed the need to have data posted in a trusted repository [1]. Numerous scientific studies are now based on publicly available remote-sensing products that seem to fit this criterion well. However, as we show here, there is a serious reproducibility issue if old versions of products become unavailable.

Recent studies have used remotely-sensed and field-measured tree cover to detect multi-modality in the tree cover frequency distribution [2–5]. If the environmental conditions such as rainfall and soil characteristics have a unimodal frequency distribution, marked modes in the frequency distribution of the state of a system may indicate the existence of alternative stable states [6]. In the case of tree cover this interpretation has important implications for understanding ecosystem responses to global climate change. Some potential caveats of interpreting remotely-sensed tree cover have been pointed out [7–9]. Here we address the fundamental issue that results may substantially change with the version of the product used. This precludes reproduction of the findings if the old version is no longer available.

For the boreal region, Scheffer et al. [3] inferred the existence of four distinct states by quantifying the modes of frequency distribution of tree cover, using the Collection 3 data of MODIS Vegetation Continuous Fields (VCF) [10]. Since then, the MODIS VCF product has been updated to Collection 5 [11]; and a Landsat VCF product of global tree cover at a fine spatial resolution of 30 meters has been developed [12]. Here we show that identification of alternative stable states in the boreal biome partly depends on the data source used, by comparing the results from these different remote-sensing products of tree cover.

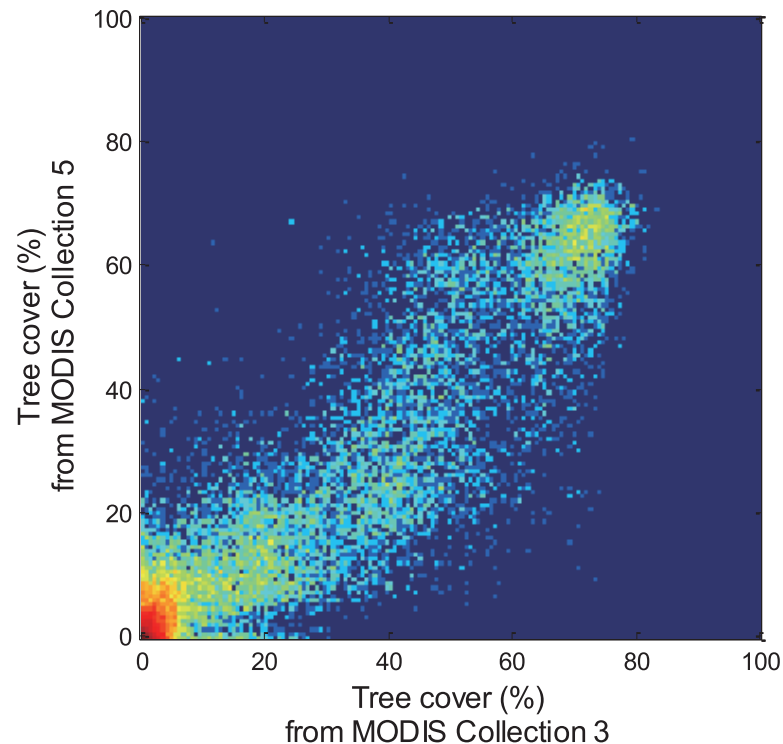
## Materials and Methods

There are major differences between these two versions of MODIS VCF data. According to the authors, new high-resolution training data and the implementation of improved data mining software have resulted in much greater accuracy in the new Collection 5 product [13]. To assess the consequence of such differences, we used the same sampling points (locations and sample dates) in the boreal region and processed the data in the same way as in our previous work [3].

## Results and Discussion

The strong correlation (Pearson's  $r = 0.861$ ,  $P < 0.001$ ,  $n = 29\,052$ ) implies general consistency between the Collections 3 and 5 (Fig 1). However, ~35% of the sampling points differ by >10% of tree cover between Collections 3 and 5. Importantly, the frequency distribution of tree cover, used as the basis for inferring multi-modality in tree cover, shows substantial discrepancy. Four modes were previously identified from the Collection 3 product, representing boreal forest (~75% tree cover), dense and sparse woodland (~45% and ~20% tree cover respectively), and a treeless state (Fig 2a). The new version of the MODIS VCF product still shows a clear multi-modal frequency distribution of tree cover. However, the modes representing sparse woodland and boreal forest now occur at a ~10% lower tree cover (Fig 2b), and the separate mode for dense woodland is not statistically significant anymore (tested by using latent class analysis performed in the *gmdistribution* function of MATLAB, S1 Fig). A closer look reveals that the Collection 3 product generally underestimates tree cover around 10–30% and 60%, while it overestimates tree cover around 40% and 70% (Fig 2c).

A mechanistic explanation of these differences would require analysis of the inherently complex calibration procedures of the VCF product [9]. However, the tri-modality suggested by the newer data seems a more robust result than the previously reported result with two (dense and sparse) woodland modes. Tri-modality is consistently detected at multiple temporal and spatial scales. During 2000–2010, annual tree cover from the MODIS Collection 5



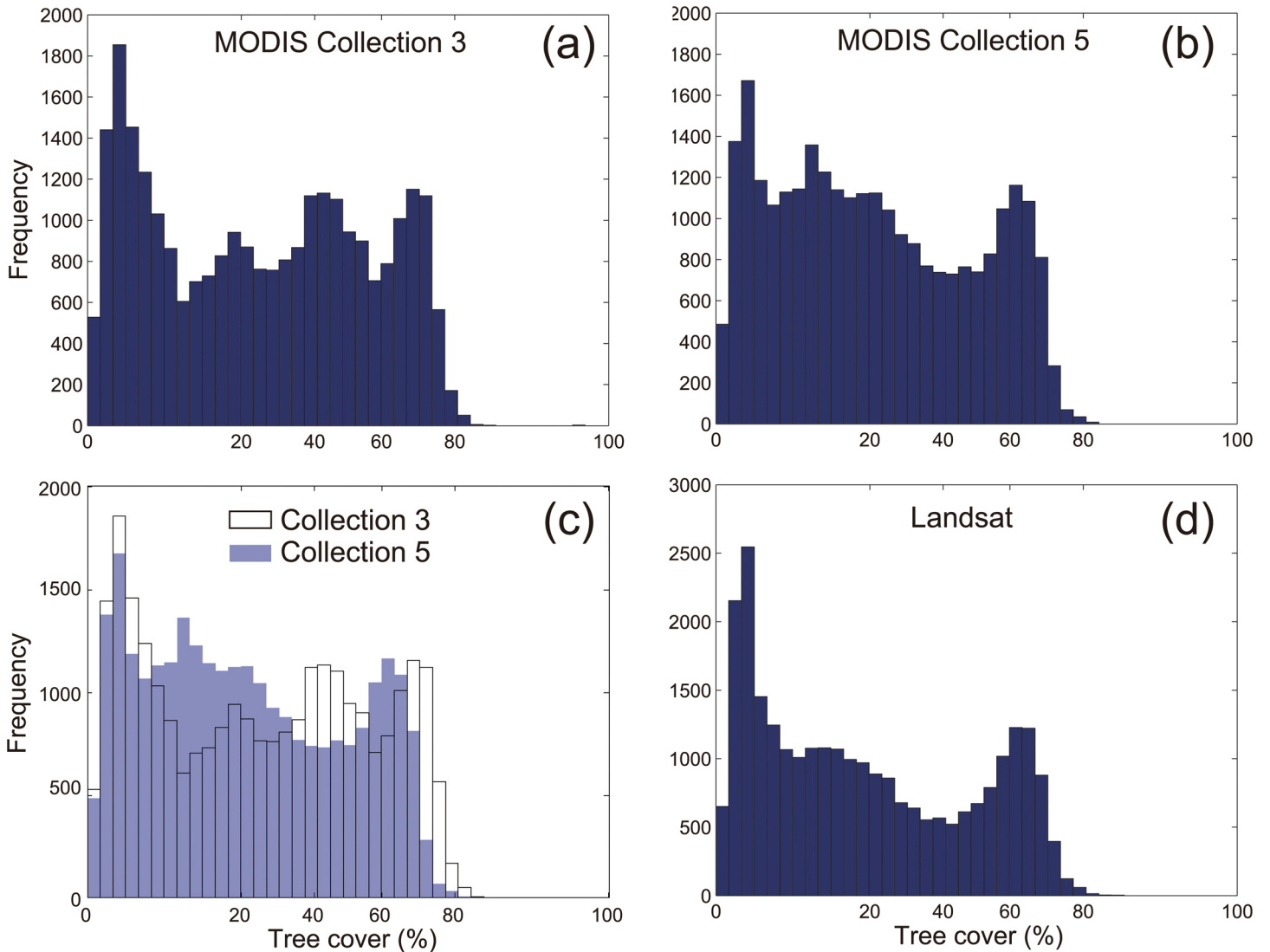
**Fig 1. Scatter density plot of tree cover in 2001 from MODIS Collection 3 against Collection 5.** Point density from low to high is indicated by the color gradient from blue to red.

doi:10.1371/journal.pone.0143014.g001

repeatedly shows three distinct modes, where the two modes representing boreal forest (~65% tree cover) and treeless state (~5% tree cover) remain highly stable; and one mode (between 15–25% tree cover) representing ‘savanna-like’ woodland is statistically significant across the 11 years (S2 Fig). Note that our analysis on these multi-temporal data aims at checking consistency of frequency distribution, rather than comparing temporal system dynamics (see [14] for the caveat of inter-annual comparison of MODIS VCF data). Being well in line with the 250-m MODIS VCF data, the 30-m resolution Landsat VCF data also detects three distinct modes around similar tree-cover values than MODIS Collection 5 (Fig 2d). Validation work has demonstrated that the Landsat VCF data with finer resolution have higher accuracy than the MODIS VCF data [12]. Particularly, the Landsat data show substantially improved performance for estimating sparse tree covers, where the MODIS data have been criticized to have greater errors [7, 15], mostly due to the limitation of spatial resolution. The consistent results from these two different remote-sensing products provide more robust evidence of tri-modal pattern of tree cover in the boreal region.

## Conclusions

Our new analysis thus suggests that, much as in the tropics [2, 16], the boreal region has three modes: one sparsely vegetated state, one distinct ‘savanna-like’ state and one forest state, which could be alternative stable states. This could reflect sensitivity to similar processes (e.g., disturbance) between the two biomes and suggests that both biomes might undergo rapid and potentially irreversible biome shifts if climate change alters the processes responsible for multi-modal patterns [2]. These results illustrate how profoundly results may change depending on the version of a remote sensing product used. Clearly, new versions of products do usually



**Fig 2. Comparison of tree-cover frequency distribution from different data sources.** MODIS Collection 3 (a) and Collection 5 (b) in 2001, frequency difference between these two versions at each tree-cover bin (c) and Landsat VCF data (d) in 2000. Tree cover percentage values have been transformed through the arcsine-squared-root transformation.

doi:10.1371/journal.pone.0143014.g002

imply an improvement. However, our findings illustrate that it is important to keep old versions available to allow independent reproduction of the results.

### Supporting Information

**S1 Fig. Comparison of fitting 1–5 normal distributions to tree cover based on the Bayesian Information Criterion (BIC).**

(PDF)

**S2 Fig. Frequency distribution of annual tree cover from MODIS Collection 5 during 2000–2010.**

(PDF)

## Author Contributions

Conceived and designed the experiments: CX M. Holmgren EHVN M. Hirota FSC MS. Performed the experiments: CX. Analyzed the data: CX. Contributed reagents/materials/analysis tools: CX EHVN M. Hirota. Wrote the paper: CX M. Holmgren EHVN M. Hirota FSC MS. Interpreted the results: CX M. Holmgren EHVN M. Hirota FSC MS.

## References

1. Nosek B, Alter G, Banks G, Borsboom D, Bowman S, Breckler S, et al. Promoting an open research culture. *Science*. 2015; 348(6242):1422–5. doi: [10.1126/science.aab2374](https://doi.org/10.1126/science.aab2374) PMID: [26113702](https://pubmed.ncbi.nlm.nih.gov/26113702/)
2. Hirota M, Holmgren M, Van Nes EH, Scheffer M. Global resilience of tropical forest and savanna to critical transitions. *Science*. 2011; 334(6053):232–5. doi: [10.1126/science.1210657](https://doi.org/10.1126/science.1210657) PMID: [21998390](https://pubmed.ncbi.nlm.nih.gov/21998390/)
3. Scheffer M, Hirota M, Holmgren M, Van Nes EH, Chapin FS. Thresholds for boreal biome transitions. *Proc Natl Acad Sci USA*. 2012; 109(52):21384–9. doi: [10.1073/pnas.1219844110](https://doi.org/10.1073/pnas.1219844110) PMID: [23236159](https://pubmed.ncbi.nlm.nih.gov/23236159/)
4. Staver AC, Archibald S, Levin SA. The global extent and determinants of savanna and forest as alternative biome states. *Science*. 2011; 334(6053):230–2. doi: [10.1126/science.1210465](https://doi.org/10.1126/science.1210465) PMID: [21998389](https://pubmed.ncbi.nlm.nih.gov/21998389/)
5. Staal A, Flores B. Sharp ecotones spark sharp ideas: comment on "Structural, physiognomic and above-ground biomass variation in savanna–forest transition zones on three continents—how different are co-occurring savanna and forest formations?" by Veenendaal et al.(2015). *Biogeosciences*. 2015; 12(18):5563–6.
6. Scheffer M, Carpenter SR. Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends Ecol Evol*. 2003; 18(12):648–56.
7. Staver AC, Hansen MC. Analysis of stable states in global savannas: is the CART pulling the horse?—a comment. *Global Ecol Biogeogr*. 2015; 24:985–7.
8. Hanan NP, Tredennick AT, Prihodko L, Bucini G, Dohn J. Analysis of stable states in global savannas—a response to Staver and Hansen. *Global Ecol Biogeogr*. 2015; 24:955–89.
9. Hanan NP, Tredennick AT, Prihodko L, Bucini G, Dohn J. Analysis of stable states in global savannas: is the CART pulling the horse? *Global Ecol Biogeogr*. 2014; 23(3):259–63.
10. Hansen M, DeFries R, Townshend J, Carroll M, Dimiceli C, Sohlberg R. Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm. *Earth Interact*. 2003; 7(10):1–15.
11. Dimiceli C, Carroll M, Sohlberg R, Huang C, Hansen M, Townshend J. Annual global automated MODIS vegetation continuous fields (MOD44B) at 250 m spatial resolution for data years beginning day 65, 2000–2010, Collection 5 percent tree cover. University of Maryland, College Park, MD, USA; 2011.
12. Sexton JO, Song X-P, Feng M, Noojipady P, Anand A, Huang C, et al. Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *International Journal of Digital Earth*. 2013; 6(5):427–48.
13. Townshend J, Hansen M, Carroll M, DiMiceli C, Sohlberg R, Huang C. User guide for the MODIS vegetation continuous fields product. Collection 5, Version 1. University of Maryland, MD, USA; 2011.
14. Carroll M, Townshend J, Hansen M, DiMiceli C, Sohlberg R, Wurster K. MODIS vegetative cover conversion and vegetation continuous fields. In: Ramachandran B, Justice CO, Abram MJ, editors. *Land Remote Sensing and Global Environmental Change*: Springer; 2011. p. 725–45.
15. Montesano P, Nelson R, Sun G, Margolis H, Kerber A, Ranson K. MODIS tree cover validation for the circumpolar taiga–tundra transition zone. *Remote Sens Environ*. 2009; 113(10):2130–41.
16. Van Nes EH, Hirota M, Holmgren M, Scheffer M. Tipping points in tropical tree cover: linking theory to data. *Global Change Biol*. 2014; 20(3):1016–21.