Introduction

- The semiarid hillslope ecosystem shows two stable states of soil depth and biomass, healthy and deserted, with a critical shift that can be caused by a change in grazing pressure.
- Although the equations for the soil and vegetation evolution in hillslopes are known at detailed resolutions, their aggregated behaviour is still poorly understood.
- Traditionally this has been explored using simplified models derived from expert knowledge, but in this study a machine learning model is used instead.

Research question

Can machine learning be used to simulate the aggregated behaviour of a hillslope ecosystem to understand the mechanisms driving desertification instead of expert-based models?

Methods

• Our knowledge of the system is captured by the time transfer function f, which describes how the state variables Z (i.e. biomass and soil depth) change over time. It can be obtained in two ways:

1. Expert knowledge

• f is defined as a set of differential equations identified through qualitative understanding of basic principles (Karssenberg et al. 2017).

2. Machine Learning

- A dataset is created by running a physically-based spatially-explicit model on a hillslope for several time steps, forcing it through changes in grazing pressure.
- The state variables Z are spatially averaged over the hillslope.
- An artificial neural network is trained to predict the rate of change from one time step to the next $(\Delta Z/\Delta t)$, capturing f.

Results

• Studying the **nullclines** ($\Delta Z/\Delta t = 0$) of the time transfer function, we can identify how the system equilibria are affected by changes in grazing pressure (Figures 1, 2).

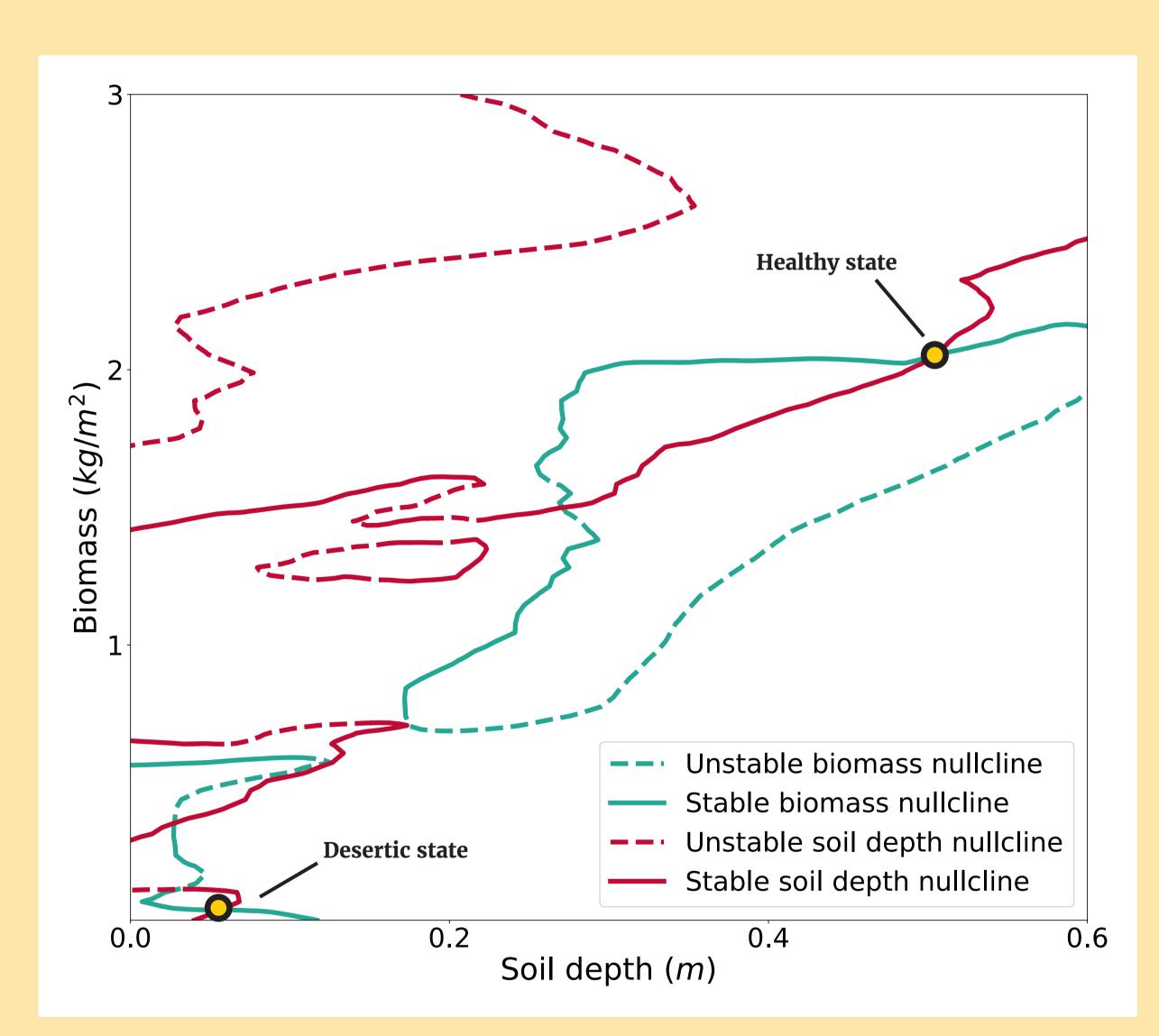


Figure 2: Stable (solid line) and unstable (dashed line) nullclines of the soil (red line) and vegetation (green line) subsystems at grazing pressure $g=1.76 \text{ kg/m}^2/\text{yr}$. The yellow circles represent the stable system equilibria.

Conclusions

- Machine learning outperforms expert-based models in estimating the system's aggregate behaviour, enabling an evidence-based knowledge of the desertification mechanisms.
- Where data is scarce or when extrapolating, the artificial neural network can become biased or even provide physically inconsistent results.
- This methodology could be extended towards other geoscientific systems in which the general behaviour is not yet well understood.

References: Karssenberg, D., Bierkens, M. F., & Rietkerk, M. (2017). Catastrophic shifts in semiarid vegetation-soil systems may unfold rapidly or slowly. The American Naturalist, 190(6), E145-E155.

Oriol Pomarol Moya¹ Derek Karssenberg¹ o.pomarolmoya@uu.nl

d.karssenberg@uu.nl 1 Department of Physical Geography, Faculty of Geosciences, Utrecht University, The Netherlands





Applying machine learning to understand hillslope desertification.

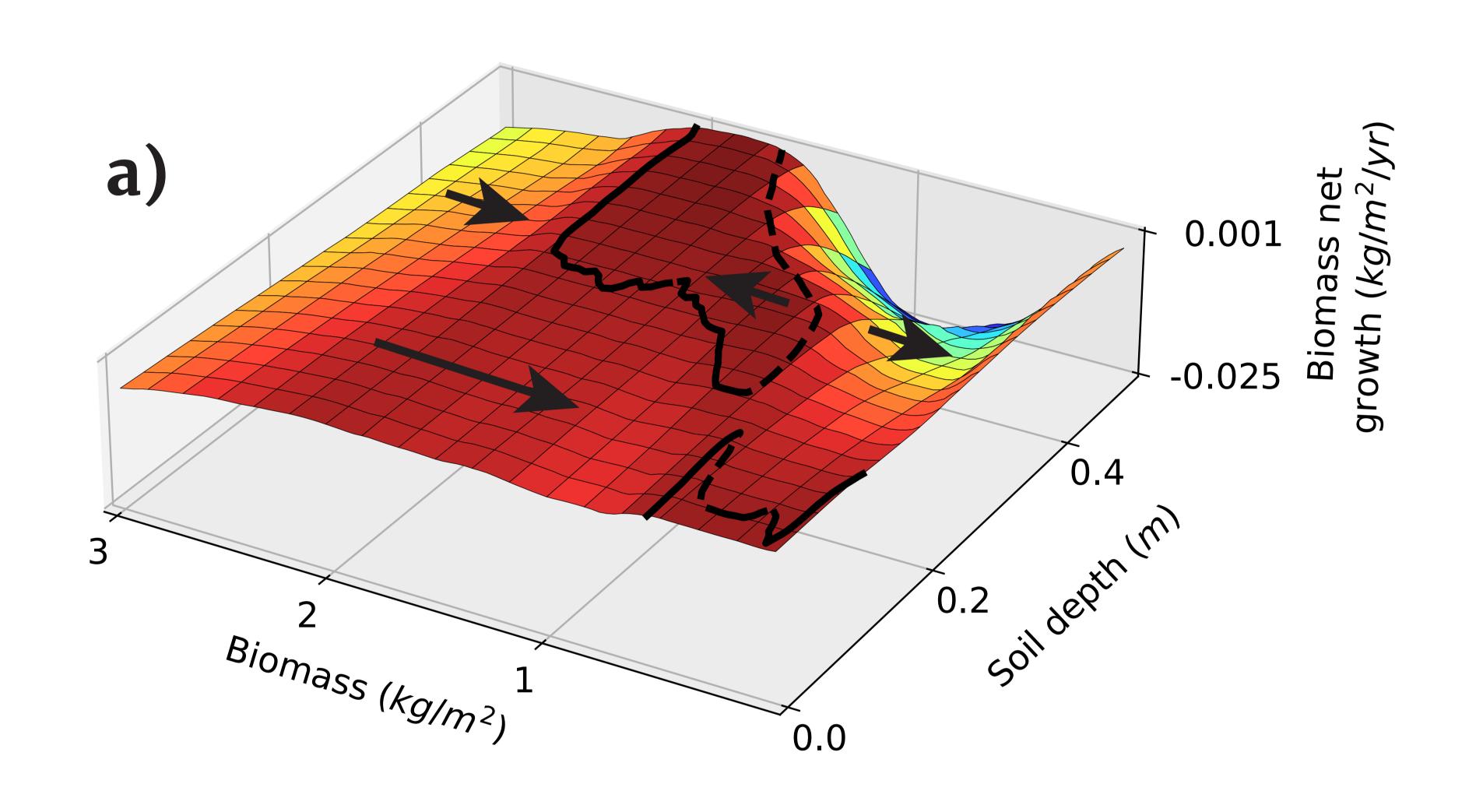
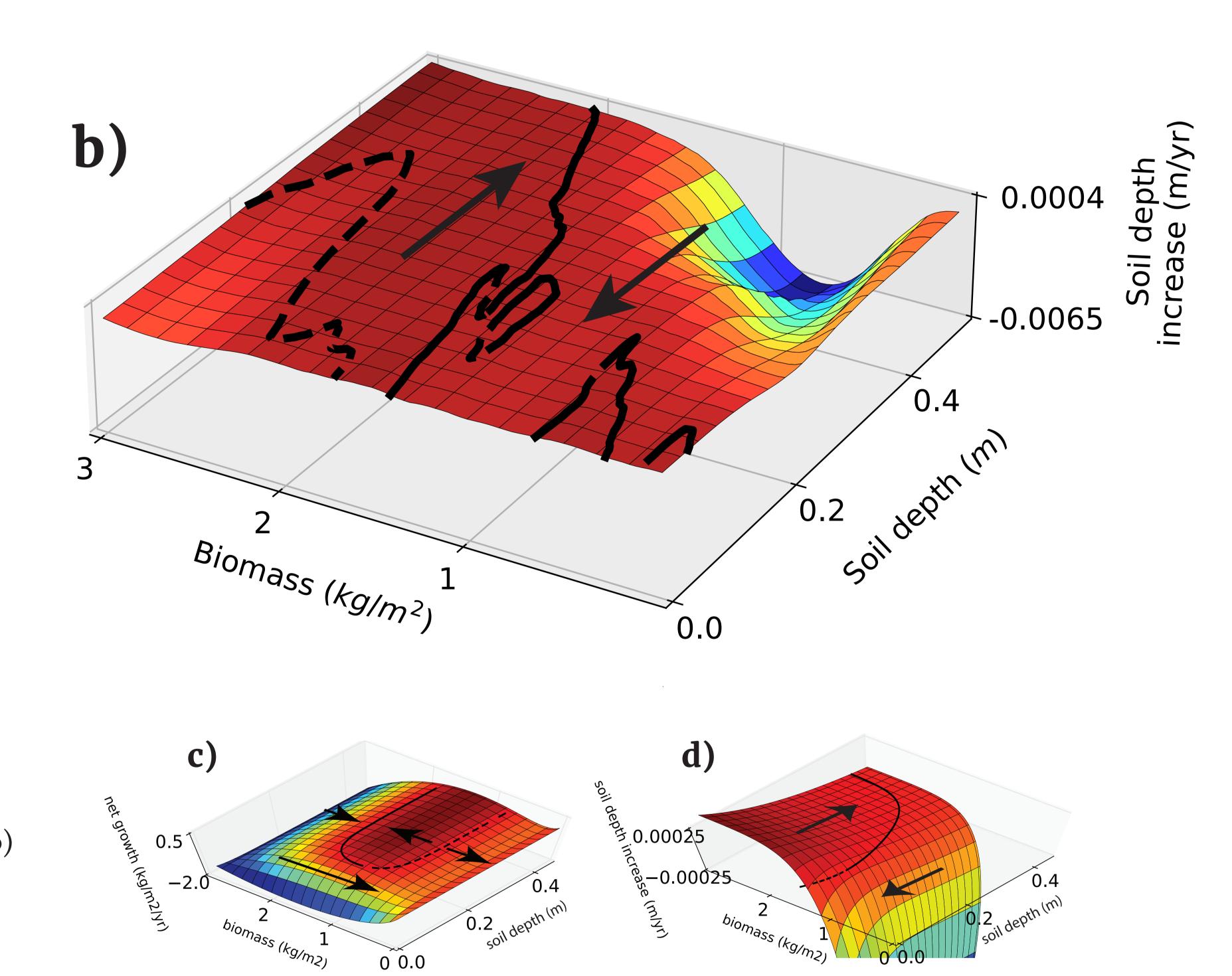


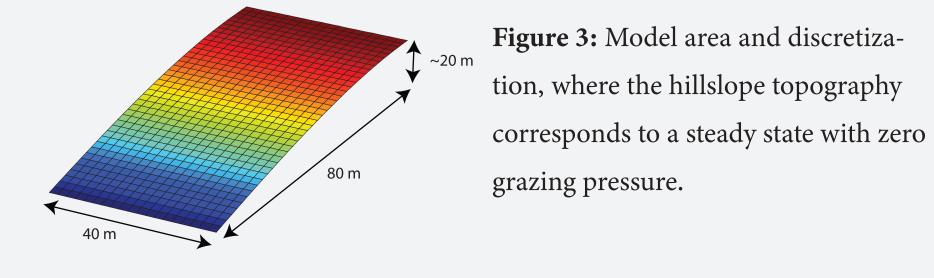
Figure 1: Biomass net growth (a,c) and soil depth increase (b,d) plotted against biomass B and soil depth D at grazing pressure $g=1.76 \ kg/m^2/yr$. Obtained using an artificial neural network (a,b) or expert-based analytical equations (c,d). Solid line, stable nullcline; dashed line, unstable nullcline. Arrows indicate the direction of change.



Additional information

Detailed model:

phological, and vegetation subsystem and the ature (Appendix A, Karssenberg et al. 2017).



Data set generation:

The model represents the hydrological, geomor- The dataset is generated running the detailed model forward in time starting at equilibrium couplings between those in a spatially distribut- for a random grazing pressure value. The system ed manner. It combines existing model formula- is forced with sinusoidal oscillations around the tions and parameters values taken from the liter- initial value, following a sharp linear increase in the end to force the transition into the desertic state to happen (Figure 4c). The spatially averaged values of biomass (B), soil depth (D) and grazing pressure (g) are stored and temporaly smoothed using a median filter.

Machine learning training:

Both fully connected artificial neural network (ANN) and random forest (RF) models have been tested using multiple hyperparameters. The input variables are B, D and g at a given time step t, while the output is the rate of change for B and D between t and $t+\Delta t$.

Model validation:

The models are validated through forward simulation of biomass and soil depth using Euler's method (Figure 4). While RF provides better estimates in forward simulation, ANNs generate more realistic time transfer functions (Figure 1).

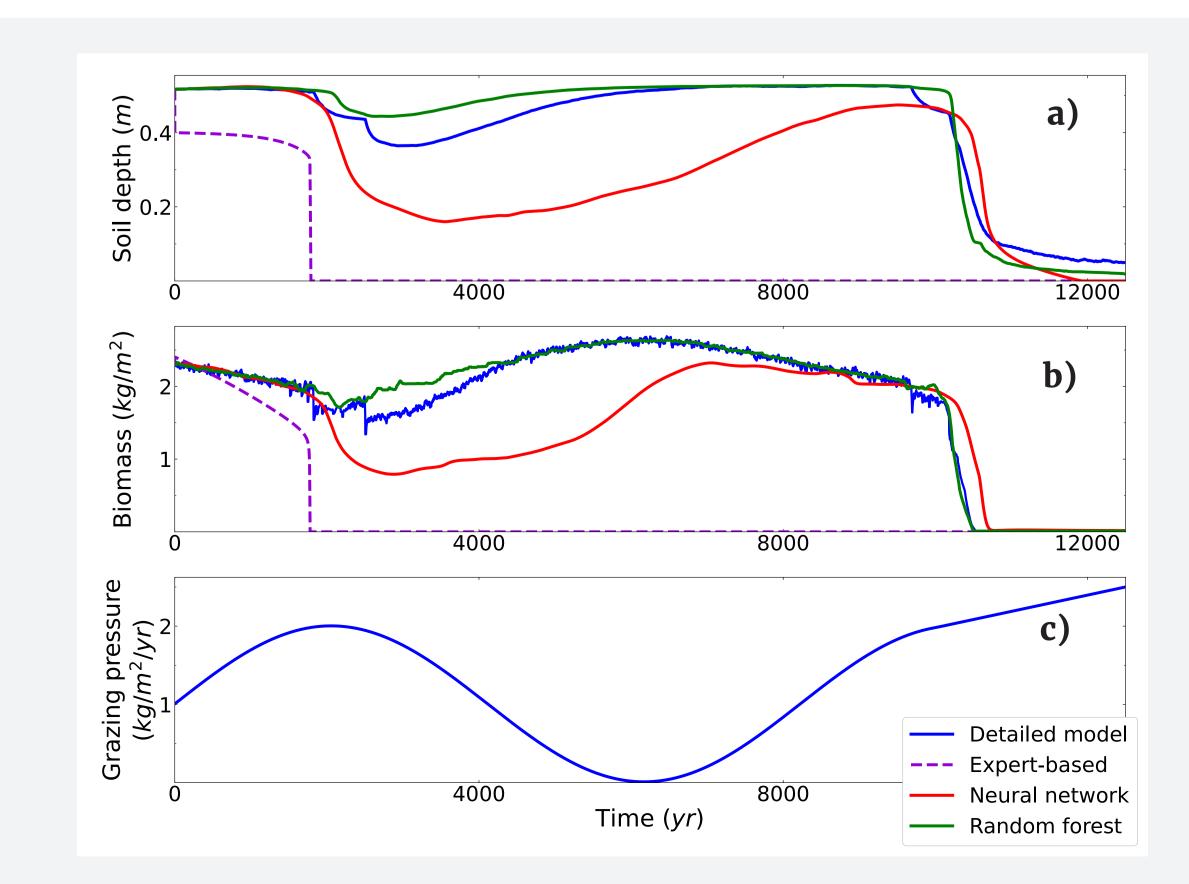


Figure 4: Forward simulation of the soil depth (a) and biomass (b) of the hillslope ecosystem forced with a sinusoidal grazing pressure pattern (c) including a linear increase at the end to induce the transition into a desertic state. The aggregated detailed model (blue line), expert-based equations (dashed violet line), artificial neural network (red line) and random forest (green line) models are compared.