# Impacts of fertilization on grassland productivity and water quality across the European Alps: insights from a mechanistic model

Martina Botter[1], Matthias Zeeman[2], Paolo Burlando[1], Simone Fatichi[3]

[1]Institute of Environmental Engineering, ETH Zurich, Zurich, Switzerland
5   [2] Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Atmospheric Environmental Research, Garmisch-Partenkirchen, Germany
[3] Department of Civil and Environmental Engineering, National University of Singapore, Singapore

*Correspondence to*: Martina Botter (botter@ifu.baug.ethz.ch)

10   **Abstract.** Alpine grasslands sustain local economy providing fodder for livestock. Intensive fertilization is common to enhance their yields, thus creating negative externalities on water quality that are difficult to evaluate without reliable estimates of nutrient fluxes. We apply a 1-D mechanistic ecosystem model, seamlessly integrating land-surface energy balance, soil hydrology, vegetation dynamics, and soil biogeochemistry aiming at assessing the grassland response to fertilization. We simulate the major water, carbon, nutrient, and energy fluxes of nine grassland plots across the broad European Alpine region.

15   We provide an unprecedent interdisciplinary model evaluation confirming its performance against observed variables from different datasets. Subsequently, we apply the model to test the influence of fertilization practices on grassland yields and nitrate ($NO_3$) losses through leaching.

Despite the generally low $NO_3$ concentration in groundwater recharge, the variability across sites is remarkable, mostly, but not exclusively, dictated by elevation. In high-Alpine sites short growing seasons lead to less efficient nitrogen (N) uptake for

20   biomass production. This combined with lower evapotranspiration rates results in higher amounts of drainage and $NO_3$ leaching to groundwater.

The local soil hydrology has a crucial role in driving the $NO_3$ use efficiency. The commonly applied fixed-threshold limit on fertilizer N input is suboptimal. We suggest that major hydrological and soil property differences across sites should be considered in the delineation of best practices or regulations for management. Using distributed maps informed with key soil

25   and climatic attributes or systematically implementing integrated ecosystem models as shown here can contribute to achieving more sustainable practices.

## 1 Introduction

Alpine grasslands are a vital resource for the European economy. They provide a variety of ecosystem services such as maintaining biodiversity, protecting soil, and offering recreational functions (Sala and Paruelo, 1997; Lamarque et al., 2011; Schirpke et al., 2017). Above all, they sustain economy providing fodder for livestock. To this purpose high yields are desired,
35  as they correspond to increasing income for farmers, hence large amounts of fertilizers are applied every year. One of the most important components of fertilizers is nitrogen. On the one hand nitrogen plays a crucial role as a nutrient contributing to the grasslands' growth. On the other hand, the surplus of nitrogen, mostly as $NO_3$, is dispersed in the environment, either through surface runoff or as leachate to groundwater. Well-known water quality issues, such as eutrophication, result directly from intensive management of grasslands (e.g., Heathwaite, 1995; Peukert et al., 2014). In order to tackle the problem, the European
40  Union (EU) approved the Nitrate Directive 91/676/EEC (EEC, 1991), which imposes limits on the yearly load of fertilizers and restricts the fertilization season. In this frame, the EU countries should pursue grassland management strategies designed with the intention of maximizing the grassland yields without damaging the environment. The task is particularly daunting given the spatial heterogeneity across the wide Alpine grassland areas. Elevation is often a major constrain for the management in Alpine environment, being low-elevation sites generally more prone than high-elevation sites to intensive management, for
45  both climatological and pragmatic reasons. Since the EU Nitrate Directive has been in effect, many studies attempted to quantify the $NO_3$ fluxes with various methods (e. g., Kronvang et al., 2009; Groenendijk et al., 2014). However, according to a questionnaire submitted to grassland experts across EU countries, Norway and Switzerland (Velthof et al., 2014), "non-academic" quantification of $NO_3$ fluxes is still mainly based on expert estimates or on prescribed values. Rarely models are applied. When they are, the most common approach is the simple empirical feed balance method, where grassland yields are
50  estimated using statistical data on feed availability for ruminants and their feed requirements. The nutrients balance is consequently estimated applying values derived from literature or direct measurements in sampled grass.

More complex modelling approaches for the quantification of nitrate losses from agricultural activity are emerging. Some European Projects have been promoted with the goal of assessing the state-of-the-art modelling tools in view of a harmonized assessment procedure across Europe. For example, the EUROHARP Project (Kronvang et al., 2009) compares available
55  distributed models applied in different European countries. The analyzed models span from conceptual models such as NLES_CAT (Simmelsgaard and Djurhuus, 1998) and MONERIS (Behrendt et al., 2003) to more complex and process-oriented models like SWAT and NL-CAT, born from the integration of ANIMO (Groenendijk et al., 2005), SWAP (Kroes and Dam, 2003), SWQN (Smit et al., 2009) and SWQL (Siderius et al., 2009). Another example is the GENESIS Project (Groenendijk et al., 2014), which compares 1-D mechanistic models integrating hydrology, vegetation dynamics, and soil biogeochemistry.
60  The performance of some models such as ARMOSA (Perego et al., 2013), CoupModel (Jansson, 2012) and EPIC (Williams et al., 1984; Sohier et al., 2009) are compared to simulated nitrate leaching rates from lysimeters. What emerges from these studies is that mechanistic models are currently not used for the purpose of supporting management because they are thought to be difficult to operate, require a large amount of input data, and require complex parameterization (see also discussions in

Fatichi et al., 2019). However, when the above limitations are overcome, mechanistic models could provide insights on
65  biomass and nutrients budget quantification that would be too difficult to obtain otherwise. They can constrain the description
of soil biogeochemical processes by means of mass and stoichiometric constraints better than empirical approaches providing
a larger predictive power ( Moorhead et al., 1996; Wieder et al., 2013; Manzoni et al., 2016).

The $NO_3$ leaching depends on the overall N cycle, which results from the interplay of various processes. Hence, mechanistic
models reproducing the whole N cycle require a tight integration with models or modules dealing with land-surface exchanges
70  of energy and water, vegetation dynamics, soil hydrological and biogeochemical processes and hydrological transport, in other
words what are called ecosystem or terrestrial biosphere models (Fatichi et al., 2016).
Different ecosystem models have been developed, but they rarely integrate all the above mentioned modules, or treat those
with a similar level of complexity. For example, models that were born as tools for the estimation of the greenhouse gases
emissions tend to simplify the soil hydrology using bucket-type approaches. Some examples of this kind of models are CERES-
75  EGC (Gabrielle et al., 1995; Gabrielle and Kengni, 1996; Hénault et al., 2005), DAYCENT ( Parton et al., 1998; 2015; Del
Grosso et al., 2000; 2002), STICS (Brisson et al., 1998; 2002; 2003) and DNDC (Li et al., 2000; 2012), although efforts have
recently been made to enhance the hydrological module of the latter (Smith et al., 2020). In other models, the soil hydrology
module is represented with much higher detail than the soil biogeochemistry module, as is the case for HYDRUS-1D/2D/3D
( Tafteh and Sepaskhah, 2012; Phogat et al., 2013) or DAISY (Hansen, 1990; 2002). More recently, ecosystem models were
80  built combining different models, each highly specialized in a specific compartment of the ecosystem. The result leads to fully-
integrated models such as ARMOSA (Perego et al., 2013), which is the combination of the hydrological model SWAP (Van
Dam, 2000), STAMINA (Ferrara et al., 2010; Richter et al., 2010) for simulating the crops dynamics, and SOILN (Bergström
et al., 1991) to represent the soil carbon and nitrogen cycles. The model SIM-STO (Pütz et al., 2018) was also created from
the merging of SIMWASER (Stenitzer, 1988) and STOTRASIM (Feichtinger, 1998), respectively focused on water fluxes
85  and nitrogen dynamics. RT-Flux-PIHM (Bao et al., 2017) combines a reactive transport (RT) model, a land-surface model
(Noah LSM, Chen et al., 2001; Shi et al., 2013) and the hydrological model PIHM (Shi et al., 2013). These models are usually
performing well when applied with targeted applications. In most of the cases they are validated only against the data
concerning one specific module that is the most important to solve the problem at hands. Hardly ever the models are
concurrently confirmed against vegetation and soil water dynamics, energy and biogeochemical fluxes. Here, we push the
90  envelope of integrated models by applying the mechanistic ecosystem model T&C-BG, which simulates the interactions among
vegetation dynamics, soil biogeochemistry, and hydrological fluxes. T&C-BG was recently developed (Fatichi et al., 2019) as
a single fully integrated model and therefore does not require many of the pragmatic simplifications necessary to externally
couple different models. We evaluate the model capabilities by applying a "common" (non-site specific) parametrization to
nine managed grassland sites across the broad alpine region, derived from previous plot-scale studies (Fatichi et al., 2014;
95  Fatichi and Pappas, 2017). First, we confirm the model against hydrological, energy, vegetation, and soil biogeochemical
observations according to the available data. Then, we run numerical experiments to estimate the grass yields and the losses

of NO₃ under different fertilization regimes and compare results across the sites. Specifically, we investigate and answer the following questions:

(1) Is a mechanistic ecosystem model provided with a "common" parameterization for Alpine managed grasslands able to reliably represent the ecosystem dynamics across 9 sites?

(2) How do grass productivity and leaching of NO₃ change in response to different fertilization scenarios across the broad Alpine region?

(3) Can mechanistic models provide guidelines for fertilization regulation?

Answering these questions, beyond allowing the assessment of a state-of-the-art ecosystem model in reproducing multiple aspects of ecosystem functioning, also provides quantitative information that can be used to enhance best grassland management practices.

## 2 Study sites and methods

### 2.1 Study sites and instrumentation

The study sites are located in European Alps and cover an altitude gradient ranging between 393 m a.s.l and 2160 m a.s.l. We used grassland sites in the Alpine region for which at least eddy covariance flux tower observations were available. For a subset of sites, additional observations about leaf area index, soil moisture, grass productivity and water and nitrogen leaching were also available. Specifically, the sites of Torgnon (TOR) and Monte Bondone (MB) in Italy, Oensingen (OEN), Chamau (CHA) and Früebüel (FRU) in Switzerland, Stubai (STU) in Austria and Fendt (FEN), Rottenbuch (ROT) and Graswang (GRA) in Germany were used. The mean annual precipitation (MAP) across sites ranges between 850 mm and 1627 mm while mean annual temperature varies between 2.9 and 9.5 °C following the elevation gradient (Table 1). All these sites are permanent grasslands ecosystems, mainly used for fodder production (Ammann et al., 2007; Hammerle et al., 2008; Vescovo and Gianelle, 2008; Kiese et al., 2018). Apart from the site TOR, which is unmanaged, the other grassland sites are managed with regular fertilizer applications and harvested with grass cuts at standard height (0.07 m).

Each site is equipped with a flux tower providing time series of micrometeorological variables, as well as energy (e.g., net radiation, latent heat, sensible heat) and carbon dioxide fluxes by means of the Eddy Covariance (EC) method (Aubinet et al., 2012). The flux tower data for all the sites but the German ones are retrieved from the FLUXNET-2015 database (Pastorello et al., 2020). German sites (FEN, ROT, GRA) belong to the Pre-Alpine Observatory of the TERENO network (Zacharias et al., 2011; Pütz et al., 2016; Kiese et al., 2018) where lysimeters are also installed in proximity of the flux-tower site (Fu et al., 2017, 2019; Zeeman et al., 2017; 2019).

For the stations FEN, ROT and GRA annual totals of lysimeter evapotranspiration (ET) and water leaching are retrieved for the years 2012-2014 from Fu et al., (2017, 2019). The same study provides annual DOC-C and nitrate-N (N-NO₃)

concentrations in the groundwater recharge. The ScaleX campaign of 2015 (Wolf et al., 2017; Zeeman et al., 2019) provided additional data concerning harvested biomass and its carbon and N content.

### 2.2 T&C-BG model

We simulate the grassland ecosystem dynamics in each site with the fully-coupled terrestrial biosphere/ecosystem model T&C-BG, introduced by Fatichi et al. (2019), who provided a complete model description. T&C-BG combines a well-tested ecohydrological model (T&C, e.g., Fatichi et al., 2012a, 2012b; Fatichi et al., 2014, 2015; Manoli et al., 2018; Mastrotheodoros et al., 2020) with a soil biogeochemistry module, which represents soil carbon and nutrient dynamics as well as plant mineral nutrition and bilateral feedbacks between nutrient availability, plant growth and soil mineralization processes. The T&C-BG model inputs include hourly flux-tower observations of microclimate (e.g., precipitation, air temperature, wind speed, relative humidity, shortwave radiation) and resolves the principal land-surface energy exchanges (e.g., net radiation, sensible and latent heat) which are interconnected with the main hydrological processes, such as evaporation and transpiration, infiltration, runoff, saturated and unsaturated zone water dynamics, groundwater recharge, as well as snow and ice hydrology dynamics. The soil hydrology module solves the 1-D Richard's equation in the vertical direction and uses a heat diffusion solution to compute the soil temperature profile. A soil-freezing module has been also recently introduced (Yu et al., 2020). The vegetation module computes photosynthesis, respiration, vegetation phenology, carbon and nutrient budget including allocation to different plant compartments and tissue turnover.

The management of vegetation such as grass cuts and fertilization can be prescribed as part of model inputs. The soil biogeochemistry module accounts for the carbon and nutrient budgets of litter and soil organic matter. Mineral nitrogen (N), phosphorous (P), and potassium (K) budgets and the nutrient leakage are also simulated. This study is particularly focused on $NO_3$. Nitrate pool in the soil depends on the net immobilization/mineralization fluxes, nitrogen uptake, nitrogen leaching, and nitrification/denitrification fluxes, with the latter that are simulated with empirical functions of the amount of $NO_3$ and environmental conditions (detailed description in Fatichi et al., 2019). $NO_3$ transport process is not solved in the soil column, but leaching is assumed to occur at the bottom of the soil column and it is proportional to the water leakage and bulk $NO_3$ concentration in soil water (Fatichi et al., 2019).

The model can be run in a distributed topographically complex domain (e.g., Mastrotheodoros et al., 2019, 2020) but here it is employed in its plot-scale version, which simply solves 1-D vertical exchanges. In summary, based solely on meteorological inputs, soil and vegetation parameters, T&C-BG simulate prognostically all the other variables, from energy, carbon, and water fluxes, to vegetation biomass, up to soil nutrient mineralization and leaching. All these processes interact with each other.

### 2.3 Model parametrization

The model allows the simulation of different vegetation types, but in this case only grass is included. In order to keep results general enough and maximize future model transferability across grassland sites, we adopt a common parametrization for all

of the nine sites with few exceptions justified by elevation dependent parameters such as threshold temperature for leaf onset.
160 Vegetation parameter selection was based on previous experience with European grasslands (Fatichi et al., 2014; Fatichi and Pappas, 2017), soil biogeochemistry parameters are currently fixed for all sites in absence of more specific information (see discussion in Fatichi et al. 2019). The site-specific soil content of clay, sand, and organic matter in each site is provided as input parameter to the model, which internally computes the hydraulic soil parameters by means of pedo-transfer functions (Saxton and Rawls 2006). To make results comparable, across all the sites we assume a soil depth equal to 1.4 m, corresponding
165 to the depth of lysimeters, which is discretized into 16 soil layers of increasing depth from the surface to the bedrock. The biogeochemistry active zone is fixed at 25 cm depth and the grass roots are assumed to be exponentially distributed with a maximum rooting depth of also 25 cm. We do not simulate the groundwater dynamics, but we compute the groundwater recharge as the water leaching at 1.40 m depth. The selected parameters are detailed in Table S1.

In each site we force the model with the local meteorological conditions, while atmospheric $CO_2$ concentrations are assumed
170 to follow the observed historical global trend (Keeling et al., 2009). Nutrients depositions in absence of local specific information for each site are set using global maps of recently observed values for nitrogen and phosphorus (Galloway et al., 2004; Mahowald et al., 2008; Vet et al., 2014).

We confirm the model performance against an ensemble of diverse variables. First, we compare simulated and observed net radiation, latent heat, sensible heat, and Gross Primary Production (GPP) using flux towers data. Second, we test the
175 hydrological module comparing the effective soil saturation (e.g., normalized soil moisture) from the model and from local measurements. Third, we confirm the vegetation module comparing the simulated biomass and/or Leaf Area Index (LAI) dynamics with data retrieved from literature wherever available (Ammann et al., 2007; 2009; Hammerle et al., 2008; Gilgen and Buchmann, 2009; Zeeman et al., 2010; Chang et al., 2013; Finger et al., 2013; Filippa et al., 2015; Prechsl et al., 2015). Finally, we compare biogeochemical fluxes in terms of harvested nitrogen and carbon, and leaching of $NO_3$ and DOC. The
180 confirmation of the soil biogeochemistry module is possible only for the sites FEN, ROT and GRA, equipped with lysimeters, even though difference in scale, soil properties, and timing of management between flux-tower footprint and lysimeters (Oberholzer et al., 2017; Mauder et al., 2018) should be accounted in the comparison.

Detailed manure input data for all the case studies were not available. To bypass this problem, whenever we did not have information about fertilization (i.e., in all sites except in the German ones), in the reference simulations we assume that the cut
185 grass is left on the ground, thus guaranteeing a nutrient application, similarly to fertilization, and a nutrient budget without major losses. While such a scenarios is unrealistic it allows to keep nutrient budgets in a dynamic equilibrium, without having to assume specific fertilization rates.

In FEN, ROT and GRA, the sites where both the flux tower and the lysimeters are co-located, grassland is managed slightly differently above the lysimeters and below the flux-tower. To validate model results we run multiple simulations, each fed
190 with the corresponding management either of flux-tower or lysimeter plots.

Since the current initial conditions of the carbon and nutrients pools in the soil are unknown, as common in modeling studies, we spin-up carbon and nutrient pools running only the soil-biogeochemistry module for 1000 years using average climatic conditions with prescribed litter inputs taken from preliminary simulations with the soil-biogeochemistry module inactive. Then we used the spun-up carbon and nutrient pools as initial conditions for the hourly-scale fully coupled simulation over the

195    period for which hourly observations are available. This last operation is repeated two times which allows reaching a dynamic equilibrium.

## 2.4 Numerical fertilization experiments

We set up numerical experiments to test the response of the study sites to different manure application regimes. First, we classify the sites based on elevation ranges and management: pre-Alpine/Intensive sites located at elevations lower than 800

200    m a.s.l. and intensively managed (CHA, OEN, FEN, ROT), Alpine/Extensive for sites with an altitude between 800 and 1000 m a.s.l. (GRA, STU, FRU) and high-Alpine/Extensive sites above 1000 m a.s.l. extensively managed, which include the two grasslands MB and TOR. It is common practice among farmers to fertilize at the beginning of the growing season and after each cut. The number of cuts and manure applications along the year decreases with elevation. Although the time intervals between cuts can be similar, the length and intensity of the winter dormant period ultimately determines growing season length

205    and, in turn, the number of cuts, the volume of locally needed organic fertilizer and the opportunity to bring that out into the field. We aligned the management strategy of our numerical experiments to these notions and used a classification of the sites into three groups. In our simulations manure is applied 6, 4, and 2 times and the grass is cut 5, 3 and 1 times, respectively, in pre-Alpine, Alpine and high-Alpine sites. Grass is cut at a height of 0.07 m, following common practice. Referring to literature (Ammann et al., 2007, 2012; Merbold et al., 2014; Fu et al., 2017) we identified a broad range of possible N yearly loads

210    applied in grasslands. For pre-Alpine sites this range spans between 50 kg ha$^{-1}$ yr$^{-1}$ (i.e., extensive management) and 500 kg ha$^{-1}$ yr$^{-1}$ (highly intensive management). This upper limit intentionally exceeds the actual management practices and allows us to analyze what can happen if fertilization loads are increased. We computed the corresponding single application amount of manure based on C:N of manure ratio equal to 8.9, which is reported for the sites of FEN, ROT and GRA (Fu et al., 2017), and is also similar to values suggested in literature (Sommerfeldt et al., 1988; Nyamangara et al., 1999) or slightly lower (Zhu,

215    2007; Kumar et al., 2010). The input of P and K is assumed to be proportional to the N input based on assigned N:P and N:K values guaranteeing non P- or K-limited conditions to the system. The specific manure amount computed for pre-Alpine sites is applied at the corresponding lower frequencies determined for Alpine and high-Alpine sites. The resulting annual loads in all the scenarios are reported in Table 2. For each fertilization scenario, in each site, we spin-up the system running the soil-biogeochemistry module for 1000 years under average climatic conditions and the specific management scenario. Thus, the

220    initial conditions of each simulation correspond to the final state of the spin-up simulation run for each specific site and management scenario.

In the result analysis, we focus our attention on the resulting N contribution to grass productivity and in the N lost as leaching at the bottom of the soil profile. The former represents one of the positive gain from agriculture as economic activity and the latter a negative externality into the environment, as most $NO_3$-leaching will ultimately reach the groundwater storage and the rivers. We evaluate the different management strategies bringing together the two indicators in an index computed as the ratio of harvested-N to N-$NO_3$ leachate. The higher the value of the index the most favorable is the management strategy for both farmers and environment. Thus, we interpret this index as a proxy of the efficiency of the grassland in profiting from the N added in fertilization. We also analyze potential relations between such efficiency index and site-specific characteristics, such as the elevation and the percentage of precipitation that becomes groundwater recharge.

## 3 Results

### 3.1 Model confirmation

The performance of the model in representing the energy and carbon fluxes measured from the flux towers is good across all the sites. The $R^2$ values of the model/observations comparison are reported in Table 3 and further goodness of fit metrics are reported in Table S2. The seasonality of the energy fluxes is well represented across all the study sites, as illustrated by the pattern of latent heat shown in Figure 2.

We show the comparison between the observed and the simulated effective saturation of the soil (Figure 3) to test the hydrological dynamics. The intra-annual pattern is generally well captured by the model. The average coefficient of determination $R^2$ equals 0.49 and the average RMSE is 0.10. Results highlight the pedo-climatic differences across the Alpine region. Some sites tend to reach field capacity quite easily (e.g., FEN, ROT and GRA), in other sites effective saturation is often below 50% (CHA, OEN, STU and FRU). The high-Alpine sites TOR and MB fall in this second case, but with pronounced peaks of saturation when snow melt occurs.

A summary of the simulated site-specific variables in terms of water, energy and carbon fluxes is shown in Table 4. The total net radiation does not vary consistently with elevation. As a result of temperature constraints, latent heat decreases with increasing elevation, leading to lower ET in high-Alpine sites, as the higher values of the Bowen ratio for the high-Alpine sites indicates. Phenology is also affected by elevation, the average day of the year when the simulated growing season starts (taken as the mean day when the biomass is higher than the biomass threshold at cut height) increases with increasing elevation. It spans from mid-March in the pre-Alpine site CHA to mid-May in the high-Alpine sites of MB and TOR. Consequently, the mean yearly GPP is higher where the growing season starts earlier. For instance, the resulting mean annual GPP in CHA is more than the double compared to TOR. However, the average GPP of the month July only, intended as a proxy for the GPP in the maximum growth period, does not differ much across the sites, and lower values are rather indicative of water limitations, testifying similar levels of productivity in the peak of the summer regardless of the site elevation.

The model confirmation against grass biomass dynamics clearly shows that snow presence on the ground limits the growing
255   season at higher altitudes (Figure 4). The model responds to the inter-annual variability of the snow cover. In years with large
snow accumulation, the growing season starts later compared to years with a less persistent snow pack. For instance, the years
2011 and 2013 in TOR are characterized by lower and higher than average snow depth respectively. The following growing
seasons is respectively anticipated and delayed in both simulations and observations.

In CHA, FRU and OEN the total simulated harvested biomass falls within, or is very close to, the range reported by
260   observational studies. Considering the large variability in published biomass estimates across sites and even within the same
site, it is difficult to conclude if such differences are a model shortcoming or simply dictated by observation uncertainty. In
OEN the LAI dynamics are also well captured. The simulated biomass and LAI patterns in FEN, ROT and GRA fit the more
detailed field data for 2015 (Figure S1). Also, the LAI in STU is simulated well, and the length of growing season and the
range of variability of available observations are matched. However, there are discrepancies on the exact dynamics of grass
265   cuts in several sites, most pronounced at STU. These are expected as grass cuts are prescribed at regular intervals in the model,
while they may occur irregularly in reality (for instance dictated by specific weather events) and they might also vary from
year to year. The simulation in TOR matches quite well the magnitude of the grass biomass and the beginning of the growing
season, but overestimates its length by approximately one month, which explains the major discrepancy between simulated
and observed LAI and leaf-biomass.

270   Simulations of the lysimeter data in FEN, ROT and GRA provide the opportunity to test the soil biogeochemical dynamics
and especially nutrient leaching as well as biomass productivity. The harvested dry matter and, consequently, also harvested
nitrogen are considerably underestimated by the model compared to the lysimeters data. The simulated mean annual leaching
of DOC in the years 2012-2014 shows values of the same magnitude of observations across all the three sites. The model
estimates accurately $N-NO_3$ leaching in ROT and GRA, while it underestimates this in FEN, where observations vary more
275   than other sites.

## 3.2 Numerical fertilization experiments

The simulated $NO_3$ concentrations in the leaching flux span two orders of magnitude ranging between 0.07 and 12 mg $L^{-1}$
(Figure 6a), compared to an environmental limit of 50 mg $L^{-1}$ in Europe (EEC, 1998). In this case we prefer comparing $NO_3$
rather than $NO_3$-N concentrations to the widely known threshold of 50 mg$NO_3$ $L^{-1}$ imposed by the European Directive. The
280   high-Alpine site TOR exhibits the highest $NO_3$ concentration, increasing almost monotonically with increasing nitrogen input
(Figure 6a). The other high-Alpine site MB shows concentrations comparable to the other sites but with higher interannual
variability. Except for TOR, in the range of N inputs between 30% and 50% of the maximum fertilization, the $NO_3$
concentration in leaching does not increase monotonically with the input (Figure 6a).

The variability of $NO_3$ leaching concentration is remarkable not only across sites belonging to different elevation classes, but
285   also within the same elevation range. The largest variability among sites of similar elevation is in the N input range of 30 to

50% of the maximum N fertilization. For example, the difference in NO$_3$ leaching concentration between STU and FRU at 50% of maximum fertilization is comparable to the difference between the leaching obtained in STU increasing the N-load from 50 to 100%. The same observation applies to the pre-Alpine sites of FEN and OEN.

Harvested N as a function of the nitrogen fertilization in the high-Alpine sites MB and TOR differs remarkably from the other
290   sites (Figure 6b). At these high-Alpine sites, the harvested-N does not considerably increase with increasing N inputs, as grass productivity is likely constrained by other factors than N. For the other sites, a clear increase of harvested N emerges for N fertilization spanning from 10% to about 40% of the maximum. For values higher than 40%, there is not much gain in simulated harvested N (and thus biomass) regardless of the increased nutrient availability.

When results are summarized in the harvested N versus N leaching space, this difference is more pronounced (Figure 6c).
295   Patterns of pre-Alpine and Alpine sites partly overlap, but occupy a distinct space away from high-Alpine sites. High-Alpine sites show a limited range of variability of harvested N compared to the range of leaching NO$_3$. Beyond a certain threshold of N input also the other pre-Alpine and Alpine sites show limited increase in harvested N. Also note that for lower N-input the NO$_3$ leaching is to be relatively stable while harvested N and thus grass biomass increase. The N contributing to grass growth is one order of magnitude higher than the leaching N in the high-Alpine sites. In comparison, for pre-Alpine and Alpine sites
300   it is two orders of magnitude higher testifying a closer N-cycle despite more intensive fertilization.

When the harvested N and the N concentration in groundwater recharge are combined in a ratio, we obtain an index, which depicts the efficiency of the fertilization practice (Figure 6d). Ideally, this index should be maximized to obtain a win-win situation which optimizes grass yield and minimizes water quality issue. When plotting the index as a function of the percentage fertilization, the pre-Alpine and Alpine sites exhibit a range where the index is maximum. This range is in between
305   20% and 60%. Specifically, CHA, OEN, FEN GRA and FRU exhibit the highest value at 20% of the fertilization input (67-100 kg N ha$^{-1}$ yr$^{-1}$), ROT at 35% of the maximum load (175 kg N ha$^{-1}$ yr$^{-1}$) and STU at 50% of the maximum load (167 kg N ha$^{-1}$ yr$^{-1}$). The high-Alpine sites do not exhibit any optimum, but only a monotonically decreasing line, as N fertilization does not stimulate growth but rather increase NO$_3$ leaching.

As a matter of fact, correlating the efficiency index with the elevation (R = -0.66 p =0.05) of each site shows a predominant
310   decreasing trend with increasing elevation (Figure 7a). However, there are exceptions such as OEN, GRA and FRU showing low index despite relatively low elevation. The correlation of the index with the fraction of precipitation, which is lost through groundwater recharge, provides an even better descriptor (R = -0.75 p =0.02) of the fertilization efficiency. We also correlate the index to the soil hydraulic conductivity (Figure S2) but the correlation is not significant (R=-0.28 p=0.47). Low water recharge fractions are associated with higher value of the N-efficiency index (CHA, FEN, ROT, STU) while sites characterized
315   by higher groundwater recharge have a low index (OEN, GRA, FRU, MB, TOR), highlighting a dominant hydrological control beyond soil-element biogeochemistry. In the sites OEN, FRU, MB and TOR the high groundwater recharge also corresponds to high soil hydraulic conductivity (Ks), while in GRA Ks is relatively small (Figure S2). On the contrary, in STU the groundwater recharge fraction is relatively small and the N-efficiency index is high despite a high hydraulic conductivity (Figure S2).

## 4 Discussion

### 4.1 Fully-integrated mechanistic ecosystem modelling: successes and limitations

While many ecosystem modeling applications have been discussed in literature, including detailed ecohydrological (e.g., Ivanov et al., 2008; Tague et al., 2013; Millar et al., 2017) and soil biogeochemistry applications (e.g., Parton et al., 1998; Kraus et al., 2014; Robertson et al., 2019), rarely, if ever, a single integrated model has been tested across different compartments and disciplines in concurrently reproducing surface energy budget, hydrological dynamics, vegetation productivity, and nitrogen budget. Here, we raise the bar to challenge T&C-BG, in reproducing these processes across nine grassland sites in the broad Alpine region. Furthermore, to ensure future model transferability and avoid local tuning, we also use the same vegetation and soil biogeochemistry parameters across all sites, with few exceptions where an elevation or latitudinal dependence of a parameter should be preserved. Despite such an "average" parameterization, the model responds surprisingly well to the challenge as energy and carbon fluxes, soil hydrology, vegetation dynamics, and $NO_3$ and DOC leaching fluxes are all with realistic magnitude and similar to observations with few notable exceptions discussed below. Also feedback between compartments are realistic in the model, as it is the case of growing season length varying depending on the date of complete snow cover disappearance or the limitations in grass growth and thus LAI at low nitrogen availability. Overall, simulations suggest that a correct representation of phenology and thus length of the growing season is a fundamental aspect of model performance as peak of the season GPP is much more similar across sites than annual GPP values (Table 4).

Despite such a positive outcome, there are a number of uncertainties in both the observations and model simulation that are relevant to highlight. It is well-known that flux towers do not close the energy budget (e.g., Foken, 1998, 2008; Wilson et al., 2002; Widmoser and Wohlfahrt, 2018; Mauder et al., 2020). The model generally overestimates the sensible heat compared to observations, thus suggesting that the missing energy is most likely attributable to sensible heat as supported by other studies (Mauder et al., 2006; Wohlfahrt et al., 2010; Liu et al., 2011) and justifies the lower $R^2$ for sensible heat compared to the other fluxes. Observations of soil water content depend on the specific soil hydraulic properties and microtopography in the location where the sensor is installed and are also often subject to temporal drifts. The model represents a vertically explicit but spatially implicit average soil moisture over the tower footprint. For this reason, even though we normalize soil moisture using effective saturation, the comparison should be seen more in qualitative terms rather than attempting to reproduce exactly the observed soil moisture values. Most important, in terms of vegetation productivity, biomass data reported from different articles present a remarkable variability despite referring to the same study site. This might depend on differences in sampling protocol and instrumentation (Zeeman et al., 2019) as well as natural spatial variability. In light of these uncertainties we do not dwell in explaining model to data differences as far as the long-term magnitude of observed biomass is similar. The only exception is the very significant underestimation of the simulated biomass compared to lysimeters measurements. We attribute large portion of this inconsistency to the well-documented lysimeter oasis/border effect, which generates crop yields and evapotranspiration fluxes 10-20% larger than larger-scale observations (Oberholzer et al., 2017). As the model captures well the pattern of biomass and carbon fluxes in the flux-tower locations, it would be difficult to justify why biomass productivity should be much different

a few hundreds of meter apart in the lysimeters under similar management (Fu et al. 2017). The model does not properly represent the inter-annual variability of harvested carbon and N, for example not capturing the higher productivity of 2013, a

355 particularly productive year due to the reduced snow cover (Zeeman et al., 2017). One explanation can be that we input to the model the same management strategy every year while local management vary from year to year. More generally, simulation results are affected by the spin-up process performed to initialize the carbon and nutrient pools in the system in absence of historical information. The induced stationarity might influence the soil organic carbon and nitrogen pools enrichment or depletion and generates discrepancies with observations. Other possible N losses in the environment such as ammonia

360 volatilization or denitrification ($N_2$) do not affect this discrepancies, being their average losses across the three sites in the order of 0.74 kgN ha$^{-1}$ yr$^{-1}$ and 0.28 kgN ha$^{-1}$ yr$^{-1}$ respectively.

Simulated $NO_3$ concentration in groundwater recharge in FEN is underestimated compared to the observed values from lysimeters. While this can be likely due to model shortcoming, the model uncertainty is not the only factor contributing to this incongruence as also the discrepancy of management between the model and reality might play a role. Moreover, the model

365 does not take into account preferential flows, which might be particularly pronounced in the lysimeters, (Schoen et al., 1999; Groh et al., 2015; Benettin et al., 2019; Shajari et al., 2019). Regardless of these existing differences, and keeping in mind that the exact value of the results is likely more uncertain that the comparison across scenarios/sites, we argue that results obtained with T&C-BG are more encouraging than discouraging and allowed us to explore the complex functioning of the overall ecosystem and to carry out virtual numerical experiments that can inform farmers and legislators.

370 **4.2 Simulated responses to fertilization and management implications**

A term of comparison for the simulated $NO_3$ leaching concentrations is the limit of 50 mg$NO_3$/L imposed on groundwater nitrate concentrations by the European Drinking Water Directive (98/83/EC, EEC, 1998). Results from the numerical experiments show that $NO_3$ leaching concentrations are generally lower or much lower than this threshold, even under highly intensive management practices.

375 The EU Nitrate Directive 91/676/EEC (EEC, 1991) imposes the maximum N fertilization rate for stable managed grasslands of 170 kgN ha$^{-1}$yr$^{-1}$. The Directive leaves room for flexibility to European countries, since they are allowed to introduce higher N fertilization loads if they demonstrate that grassland absorbs it efficiently. However, if the monitoring of surface and groundwater $NO_3$ concentration reveal water quality issues, then EU can intervene with an infringement proceeding. As an example, Germany in 2017 re-adapted the national Fertilizer Ordinance of 2007 after an infringement proceeding by the EU

380 (Kuhn, 2017).

The threshold of 170 kgN ha$^{-1}$yr$^{-1}$ is computed on the basis of a $NO_3$ input-output balance at the farm scale. In such a simple computation the $NO_3$ losses in the environment are assumed to be 30% of the input (IPCC, 2000, 2006). These losses include both groundwater recharge and surface runoff. Although some countries have lowered this threshold to 10%, such as Ireland, or to 20%, such as Switzerland, there is evidence that the threshold is quite high compared to observed losses in grasslands

385 (Eder et al., 2015). Our study confirms such a finding highlighting NO₃ concentrations in the order of 0.1-12 mg L⁻¹, although
we only consider the NO₃ losses with groundwater recharge and neglect those through surface runoff. The surface runoff
contribution to NO₃ losses is more likely to be higher in the high-Alpine sites, where the winter snowpack melting on a frozen
soil favors surface runoff. However, there is evidence that NO₃ losses through surface runoff are usually small compared to
losses through groundwater recharge ( Jackson et al., 1973; Casson et al., 2008). Despite the noticeable inter-site variability,

390 the simulated losses of NO₃ into groundwater are generally lower than 10% of the N-input, even under highly intensive
fertilization scenarios (Figure 6c).

Among the N inputs explored in the analysis, the value which guarantees the maximization of the N-efficiency ratio is quite
close or lower than the limit of 170 kgN ha⁻¹ yr⁻¹ in pre-Alpine and Alpine sites (Figure 6d). Thus, we confirm that such a limit
is a reasonable upper threshold. However, an emerging aspect from the simulations is the large variability across sites in NO₃

395 leaching and response to fertilization. While model uncertainties exist and we are unlikely to capture the exact magnitude of
all N-fluxes at all sites, this variability is likely underestimated rather than overestimated by the model due to simplifying
assumptions and commonality of parameters. The NO₃ leaching and the grassland yields depend on the degree at which
nitrogen is limiting productivity versus other environmental factors (e.g., temperature), on the capability of vegetation in
uptaking nutrients, and on the hydrological characteristics of the sites. A soil favoring water drainage or a fast snowmelt at the

400 end of the winter generates larger amount of N leaching into groundwater. Most likely one factor does not exclude the other
(Figure 7 and Figure S2), but the combination of both soil properties and hydrological regime drives the correlation between
N leaching concentration and groundwater recharge fraction. The N uptake efficiency of grassland depends more on the
nutrient demand with a longer growing season (Zeeman et al. 2017) higher temperatures and lack of water stress favoring the
N demand, even though the importance of the different aspects are difficult to disentangle (Lü et al., 2014). Literature shows

405 that also the richness of species composing the grassland might favor nutrient uptake (Tilman et al., 1996; Spehn et al., 2005;
Niklaus et al., 2006), but this type of ecological feedback is not implemented in the model, which simply considers an "average
species composition" as a representative grassland. In the simulations, the length of the growing season at OEN, GRA and
FRU is comparable to the other pre-Alpine and Alpine sites (Table 4). However, simulated NO₃ leaching is higher and the
produced biomass is lower. This is especially true at low N fertilization levels. This further remarks the important role of local

410 hydrology (Fig 6a-c) on N-cycle.

As local hydrological and soil conditions (and potentially also different site biogeochemical history and species composition–
not considered here) modify the grass response to N fertilization, disparities among farmers might emerge simply as a function
of location. Farmers owing a field in an area characterized by higher groundwater recharge might be disadvantaged because
losses of NO₃ to groundwater could be higher and yields limited for the same amount of N fertilization. On the contrary,

415 farmers owing fields in hydrological favorable sites might be advantaged. Such disparities should be taken into account in two
ways: first in delineating location specific fertilization limits and second in applying some sort of compensation for
locations/regions that are disadvantaged. A good practice in this direction is the proof of ecological performance (PEP),

introduced in Switzerland (Swiss Federal Council, 1998). Direct payments are transferred to farmers who join this program, which imposes a wealth of rules in order to preserve the environment. Among the requirements the computation of the farm

420    $NO_3$ balance needs to be also provided.

On the basis of modeling results, regulations based on a fixed-threshold for the whole European grasslands could be largely suboptimal. Even if we might underestimate losses because of model uncertainties and because we do not consider surface runoff, resulting $NO_3$ losses are abundantly lower than 30% of the N-input.

Overall, there is still leverage for enhancing N fertilization before reaching the threshold assumed by the EU Directive.

425    However, we also showed that higher N-load do not necessarily correspond to higher yields, especially above certain thresholds. Therefore, even though $NO_3$ leaching is low, an effective increase of manure application might not be beneficial for grass yield and should be evaluated on the basis of the site-specific characteristics.

We suggest as a possible alternative and better strategy the definition of thresholds on fertilizer input based on a distributed mapping of the landscape. In absence of more precise information, agricultural areas could be classified based on the crop $NO_3$

430    use efficiency, soil type, and hydrological characteristics. For instance, Klammler et al. (2013) provide and interesting example of $NO_3$ leaching mapping in a case study in Austria. Alternatively, mechanistic ecosystem models as used in this study could be employed to identify ranges of optimal fertilization levels for different sites, groups of sites, or even in a fully distributed manner. This process could be demanding in terms of resources as requires advanced expertise, sufficient data to constrain the models, and adequate modelling tools (Decrem et al., 2007) . However, with increasing computational capabilities and data

435    availability to constrain model parameters, it is likely that mechanistic modeling approaches might become more popular and an essential tool for fostering the mapping process and provide distributed information to refine environmental regulations.

## 5 Conclusion

We simulated the dynamics of nine managed grassland sites across the broad European Alpine region. We applied the mechanistic model T&C-BG which fully integrates land-surface mass and energy fluxes, soil hydrology, vegetation dynamics,

440    and soil biogeochemistry. The model was confirmed to reproduce realistic magnitude and temporal dynamics of ecohydrological variables across multiple compartments in an effort of model evaluation that goes beyond many of the existing attempts to confirm ecosystem models. The model was subsequently used to quantify the impact of different N fertilization scenarios, with focus on grass productivity and $NO_3$ concentration in groundwater recharge. Simulations reveal that although groundwater recharge concentrations are relatively small and well below environmental policy limits, there is high variability

445    across grasslands, also for sites located at similar elevation. Such variability is mainly driven by local environmental controls on productivity that reflects in the grass capability to uptake nutrients and by the differences in the hydrological regime summarized as the fraction of precipitation that becomes recharge. We suggest that these factors should be taken into account by legislators while defining thresholds on fertilizer loads. Guidelines based on the site-specific rather than based on fixed

thresholds across large regions would favor the maximization of grass yields while allowing preserving similar water quality

450 targets. Fully-integrated mechanistic ecosystem models as employed here have a big potential as tools to construct these maps under current and future climate scenarios.

**Code availability**

The model code is available at the link: https://hyd.ifu.ethz.ch/research-data-models/t-c.html (last access July 2020).

**Data availability**

455 FLUXNET data for running and confirming T&C-BG plot-scale simulations were downloaded from https://fluxnet.fluxdata.org/. Specifically we downloaded data for the stations Chamau (10.18140/FLX/1440131), Oensingen (10.18140/FLX/1440135), Früebüel (10.18140/FLX/1440133), Stubai (10.18140/FLX/1440121), Monte Bondone (10.18140/FLX/1440170) and Torgnon (10.18140/FLX/1440237). Data for the sites Fendt, Rottenbuch and Graswang are available from the TERENO data discovery portal (https://www.tereno.net).

460 **Author contribution**

MB and SF conceived the ideas and discussions with MZ supported their development. MB and SF designed the experiments, MB ran simulations and developed the analyses. All authors contributed to the writing.

**Competing interests**

The authors declare that they have no conflict of interest.

465 **Acknowledgements**

This study was supported by the DAFNE project (https://dafne.ethz.ch/), funded by the Horizon 2020 programme WATER 2020 of the European Union (grant agreement no. 690268).

The authors acknowledge Dr. Gianluca Filippa from the Environmental Protection Agency of Aosta Valley for sharing data of the site TOR. A special acknowledgment to Dr. Ralf Kiese from KIT-IfU for the insightful comments to the manuscript.

470 This study built upon TERENO/ScaleX cooperation and we thank Benjamin Wolf, Nadine Ruehr, Heather Shupe, Martina Bauerfeind, Carsten Malchow, Maximilian Graf, Matthias Mauder (all KIT/IMK-IFU) and the Scientific Team of ScaleX Campaign 2015 for their contribution. The TERrestrial Environmental Observatory (TERENO) infrastructure is funded by the

ATMO program of the Helmholtz Association and the Federal Ministry of Education and Research. MZ received support from the German Research Foundation (DFG; project ZE 1006/2-1).

475 **References**

Ammann, C., Flechard, C. R., Leifeld, J., Neftel, A., and Fuhrer, J.: The carbon budget of newly established temperate grassland depends on management intensity, Agriculture, Ecosystems and Environment, 121(1–2), 5–20. https://doi.org/10.1016/j.agee.2006.12.002, 2007.

Ammann, C., Wolff, V., Marx, O., Brümmer, C., and Neftel, A.: Measuring the biosphere-atmosphere exchange of total
480        reactive nitrogen by eddy covariance, Biogeosciences, 9, 4247–4261, 2012.

Ammann, Christof, Spirig, C., Leifeld, J., and Neftel, A.: Assessment of the nitrogen and carbon budget of two managed temperate grassland fields. Agriculture, Ecosystems and Environment, 133(3–4), 150–162, https://doi.org/10.1016/j.agee.2009.05.006, 2009.

Aubinet, M., Vesala, T., and Papale, D.: Eddy Covariance: A Practical Guide to Measurement and Data Analysis, Springer,
485        2012.

Bao, C., Li, L., Shi, Y., and Duffy, C.: Understanding watershed hydrogeochemistry: 1. Development of RT-Flux-PIHM, Water Resources Research, *53*(3), 2328–2345, https://doi.org/10.1002/2016WR018934, 2017.

Behrendt, H., Bach, M., Kunkel, R., Opitz, D., Pagenkopf, W. G., Scholz, G., and Wendland, F.: Nutrient Emissions into River Basins of Germany on the Basis of a Harmonized Procedure, http://www.umweltbundesamt.de, 2003.

490 Benettin, P., Queloz, P., Bensimon, M., McDonnell, J. J., and Rinaldo, A.: Velocities, Residence Times, Tracer Breakthroughs in a Vegetated Lysimeter: A Multitracer Experiment, Water Resources Research, 55(1), 21–33. https://doi.org/10.1029/2018WR023894, 2019

Bergström, L., Johnsson, H., and Torstensson, G.: Simulation of soil nitrogen dynamics using the SOILN model, Fertilizer Research, *27*(2–3), 181–188, https://doi.org/10.1007/BF01051126, 1991.

495 Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y. M., Cellier, P., Debaeke, P., Gaudillère, J. P., Hénault, C., Maraux, F., Seguin, B., and Sinoquet, H.: An overview of the crop model STICS, European Journal of Agronomy, *18*(3–4), 309–332. https://doi.org/10.1016/S1161-0301(02)00110-7, 2003.

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M. H., Ruget, F., Nicoullaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R.,
500        Durr, C., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J.-M., Meynard, J. M., and Delécolle, R.: STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. Agronomie, 18(5–6), 311–346. https://doi.org/10.1051/agro:19980501, 1998.

Brisson, N., Ruget, F., Gate, P., Lorgeou, J., Nicoullaud, B., Tayot, X., Plenet, D., Jeuffroy, M.-H., Bouthier, A., Ripoche, D.,

505   Mary, B., and Justes, E.: STICS: a generic model for simulating crops and their water and nitrogen balances. II. Model validation for wheat and maize, Agronomie, 22(1), 69–92. https://doi.org/10.1051/agro:2001005, 2002.

Casson, J.P., Olson, B.M., Little, J.L., and Nolan, S.C.: Assessment of Environmental Sustainability in Alberta's Agricultural Watersheds Project. Volume 4: Nitrogen loss in surface runoff. Alberta Agriculture and Rural Development, Lethbridge, Alberta, Canada. 71 pp, 2008.

510   Chang, J. F., Viovy, N., Vuichard, N., Ciais, P., Wang, T., Cozic, A., Lardy, R., Graux, A. I., Klumpp, K., Martin, R., and Soussana, J. F.: Incorporating grassland management in ORCHIDEE: model description and evaluation at 11 eddy-covariance sites in Europe. Geoscientific Model Development, 2013.

Chen, F., and J. Dudhia: Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity, Mon. Wea. Rev., **129**, 569–585, https://doi.org/10.1175/1520-
515   0493(2001)129<0569:CAALSH>2.0.CO;2, 2001.

Decrem, M., Spiess, E., Richner, W., and Herzog, F.: Impact of Swiss agricultural policies on nitrate leaching from arable land, Agron. Sustain. Dev, *27*, 243–253. https://doi.org/10.1051/agro:2007012, 2007.

Del Grosso, S. J., Parton, W. J., Mosier, A. R., Ojima, D. S., Kulmala, A. E., and Phongpan, S.: General model for N2O and N2 gas emissions from soils due to dentrification, Global Biogeochemical Cycles, 14(4), 1045–1060.
520   https://doi.org/10.1029/1999GB001225, 2000.

Del Grosso, S., Ojima, D., Parton, W., Mosier, A., Peterson, G., and Schimel, D.: Simulated effects of dryland cropping intensification on soil organic matter and greenhouse gas exchanges using the DAYCENT ecosystem model, Environmental Pollution, 116(SUPPL. 1), S75–S83. https://doi.org/10.1016/S0269-7491(01)00260-3, 2002.

Eder, A., Blöschl, G., Feichtinger, F., Herndl, M., Klammler, G., Hösch, J., Erhart, E., and Strauss, P.: Indirect nitrogen losses
525   of managed soils contributing to greenhouse emissions of agricultural areas in Austria: results from lysimeter studies, Nutrient Cycling in Agroecosystems, 101(3), 351–364. https://doi.org/10.1007/s10705-015-9682-9, 2015.

EEC.: Council Directive 19/676/EEC of 12 December, 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources. (375), 1, 1991.

EEC.: Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption, 1998.

530   Fatichi, S., Ivanov, V. Y., and Caporali, E.: A mechanistic ecohydrological model to investigate complex interactions in cold and warm water-controlled environments: 1. Theoretical framework and plot-scale analysis. *Journal of Advances in Modeling Earth Systems*, *4*(5). https://doi.org/10.1029/2011MS000086, 2012a.

Fatichi, S., Ivanov, V. Y., and Caporali, E.: A mechanistic ecohydrological model to investigate complex interactions in cold and warm water-controlled environments: 2. Spatiotemporal analyses. *Journal of Advances in Modeling Earth Systems*,
535   *4*(5). https://doi.org/10.1029/2011MS000087, 2012b.

Fatichi, S., Katul, G. G., Ivanov, V. Y., Pappas, C., Paschalis, A., Consolo, A., Kim, J., and Burlando, P.: Abiotic and biotic controls of soil moisture spatiotemporal variability and the occurrence of hysteresis. *Water Resources Research*, *51*(5), 3505–3524. https://doi.org/10.1002/2014WR016102, 2015.

Fatichi, S., Manzoni, S., Or, D., and Paschalis, A.: A Mechanistic Model of Microbially Mediated Soil Biogeochemical Processes: A Reality Check. Global Biogeochemical Cycles, *33*(6), 2018GB006077. https://doi.org/10.1029/2018GB006077, 2019.

Fatichi, S., and Pappas, C.: Constrained variability of modeled T:ET ratio across biomes. Geophysical Research Letters, 44(13), 6795–6803. https://doi.org/10.1002/2017GL074041, 2017.

Fatichi, S., Pappas, C., and Ivanov, V. Y.: Modeling plant-water interactions: an ecohydrological overview from the cell to the global scale, Wiley Interdisciplinary Reviews: Water, 3(3), 327–368. https://doi.org/10.1002/wat2.1125, 2016.

Fatichi, S., Zeeman, M. J., Fuhrer, J., and Burlando, P.: Ecohydrological effects of management on subalpine grasslands: From local to catchment scale, Water Resources Research, 50(1), 148–164. https://doi.org/10.1002/2013WR014535, 2014.

Feichtinger, F.: *STOTRASIM*–Ein Modell zur Simulation der Stickstoffdynamik in der ungesättigten Zone eines Ackerstandortes, 1998.

Ferrara, R. M., Trevisiol, P., Acutis, M., Rana, G., Richter, G. M., and Baggaley, N.: Topographic impacts on wheat yields under climate change: Two contrasted case studies in Europe. Theoretical and Applied Climatology, 99(1–2), 53–65. https://doi.org/10.1007/s00704-009-0126-9, 2010.

Filippa, G., Cremonese, E., Galvagno, M., Migliavacca, M., Morra di Cella, U., Petey, M., and Siniscalco, C.: Five years of phenological monitoring in a mountain grassland: inter-annual patterns and evaluation of the sampling protocol, International Journal of Biometeorology, 59(12), 1927–1937, https://doi.org/10.1007/s00484-015-0999-5, 2015.

Finger, R., Gilgen, A. K., Prechsl, U. E., and Buchmann, N.: An economic assessment of drought effects on three grassland systems in Switzerland, Regional Environmental Change, 13(2), 365–374, https://doi.org/10.1007/s10113-012-0346-x, 2013.

Foken, T.: Die scheinbar ungeschlossene Energiebilanz am Erdboden - eine Herausforderung an die Experimentelle Meteorologie. Sitzungsberichte der Leibniz-Sozietät, 1998.

Foken, T.: The energy balance closure problem: an overview. Ecological Applications, 18(6), 1351–1367. https://doi.org/10.1890/06-0922.1, 2008.

Fu, J., Gasche, R., Wang, N., Lu, H., Butterbach-Bahl, K., and Kiese, R.: Impacts of climate and management on water balance and nitrogen leaching from montane grassland soils of S-Germany, Environmental Pollution, 229, 119–131. https://doi.org/10.1016/J.ENVPOL.2017.05.071, 2017.

Fu, J., Gasche, R., Wang, N., Lu, H., Butterbach-Bahl, K., and Kiese, R. : Dissolved organic carbon leaching from montane grasslands under contrasting climate, soil and management conditions, Biogeochemistry, 145(1–2), 47–61. https://doi.org/10.1007/s10533-019-00589-y, 2019.

Gabrielle, B., and Kengni, L., Analysis and Field-Evaluation of the CERES Models' Soil Components: Nitrogen Transfer and Transformations. Soil Science Society of America Journal, 60(1), 142–149. https://doi.org/10.2136/sssaj1996.03615995006000010023x, 1996.

Gabrielle, B., Menasseri, S., and Houot, S.: Analysis and Field Evaluation of the Ceres Models Water Balance Component.

Soil Science Society of America Journal, 59(5), 1403–1412. https://doi.org/10.2136/sssaj1995.03615995005900050029x, 1995.

575 Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R., and Vörösmarty, C. J., Nitrogen cycles: Past, present, and future, Biogeochemistry, 70(2), 153–226, https://doi.org/10.1007/s10533-004-0370-0, 2004.

Galvagno, M., Wohlfahrt, G., Cremonese, E., Rossini, M., Colombo, R., Filippa, G., Julitta, T., Manca, G., Siniscalco, C.,
580 Morra di Cella, U., and Migliavacca, M.: Phenology and carbon dioxide source/sink strength of a subalpine grassland in response to an exceptionally short snow season. Environmental Research Letters, 8(2). https://doi.org/10.1088/1748-9326/8/2/025008, 2013.

Gianelle, D., Vescovo, L., Marcolla, B., Manca, G., and Cescatti, A.: Ecosystem carbon fluxes and canopy spectral reflectance of a mountain meadow. International Journal of Remote Sensing, 30(2), 435–449.
585 https://doi.org/10.1080/01431160802314855, 2009.

Gilgen, A. K., and Buchmann, N.: Response of temperate grasslands at different altitudes to simulated summer drought differed but scaled with annual precipitation. Biogeosciences, 6(11), 2525–2539. https://doi.org/10.5194/bg-6-2525-2009, 2009.

Gilmanov, T. G., Soussana, J. F., Aires, L., Allard, V., Ammann, C., Balzarolo, M., Barcza, Z., Bernhofer, C., Campbell, C. L., Cernusca, A., Cescatti, A., Clifton-Brown, J., Dirks, B. O. M., Dore, S., Eugster, W., Fuhrer, J., Gimeno, C.,
590 Gruenwald, T., Haszpra, L., …, and Wohlfahrt, G.: Partitioning European grassland net ecosystem CO2 exchange into gross primary productivity and ecosystem respiration using light response function analysis, Agriculture, Ecosystems and Environment, 121(1–2), 93–120, https://doi.org/10.1016/j.agee.2006.12.008, 2007.

Groenendijk, P., Renaud, L. V., and Roelsma, J.: Prediction of nitrogen and phosphorus leaching to groundwater and surface waters; process descriptions of the animo4.0 model. Alterra, 2005.

595 Groenendijk, P., Heinen, M., Klammler, G., Fank, J., Kupfersberger, H., Pisinaras, V., Gemitzi, A., Peña-Haro, S., García-Prats, A., Pulido-Velazquez, M., Perego, A., Acutis, M., and Trevisan, M.: Performance assessment of nitrate leaching models for highly vulnerable soils used in low-input farming based on lysimeter data. Science of the Total Environment, 499, 463–480. https://doi.org/10.1016/j.scitotenv.2014.07.002, 2014.

Groh, J., Pütz, T., Jülich, F., Vanderborght, J., and Vereecken, H.: Estimation of evapotranspiration and crop coefficient of an
600 intensively managed grassland ecosystem with lysimeter measurements, 16. Gumpensteiner Lysimetertagung 2015, 107–112. https://www.researchgate.net/publication/275533480, 2015.

Hammerle, A., Haslwanter, A., Tappeiner, U., Cernusca, A., and Wohlfahrt, G.: Leaf area controls on energy partitioning of a temperate mountain grassland. Biogeosciences, 5(2), 421–431, https://doi.org/10.5194/bg-5-421-2008, 2008.

Hansen, S.: Equation Section 1 Daisy, a flexible Soil-Plant-Atmosphere system Model, 2002.

605 Hansen, S., Jensen, H. E., Nielsen, N. E., and Svendsen, H.: DAISY: Soil plant atmosphere system model, 1990.

Heathwaite, L.: Sources of eutrophication: hydrological pathways of catchment nutrient export. In Man's Influence on

Freshwater Ecosystems and Water Use (Issue 230). IAHS Publ., 1995.

Hénault, C., Bizouard, F., Laville, P., Gabrielle, B., Nicoullaud, B., Germon, J. C., and Cellier, P.: Predicting in situ soil N2O emission using NOE algorithm and soil database. Global Change Biology, 11(1), 115–127. https://doi.org/10.1111/j.1365-2486.2004.00879.x, 2005.

IPCC.: IPCC - Task Force on National Greenhouse Gas Inventories. https://www.ipcc-nggip.iges.or.jp/public/gp/english/, 2000.

IPCC.: IPCC - Overview 2 2006 IPCC Guidelines for National Greenhouse Gas Inventories. http://www.ipcc-nggip.iges.or.jp/, 2006.

Ivanov, V. Y., Bras, R. L., and Vivoni, E. R.: Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks, Water Resources Research, 44(3), https://doi.org/10.1029/2006WR005588, 2008.

Jackson, W. A., Asmussen, L. E., Hauser, E. W., and White, A. W.: Nitrate in Surface and Subsurface Flow from a Small Agricultural Watershed. Journal of Environmental Quality, 2(4), 480–482. https://doi.org/10.2134/jeq1973.00472425000200040017x, 1973.

Jansson, P. E.: CoupModel: Model Use, Calibration, and Validation, Transactions of the ASABE, 55(4), 1337–1346. https://doi.org/10.13031/2013.42245, 2012.

Keeling, R. F., Piper, S. C., Bollenbacher, A. F., and Walker, J. S.: Atmospheric CO2 records from sites in the sio air sampling network, in trends: A compendium of data on global change, 2009.

Kiese, R., Fersch, B., Baessler, C., Brosy, C., Butterbach-Bahl, K., Chwala, C., Dannenmann, M., Fu, J., Gasche, R., Grote, R., Jahn, C., Klatt, J., Kunstmann, H., Mauder, M., Rödiger, T., Smiatek, G., Soltani, M., Steinbrecher, R., Völksch, I., …, and Schmid, H. P.: The TERENO Pre-Alpine Observatory: Integrating Meteorological, Hydrological, and Biogeochemical Measurements and Modeling, Vadose Zone Journal, 17(1), 0. https://doi.org/10.2136/vzj2018.03.0060, 2018.

Klammler, G., Kupfersberger, H., Rock, G., and Fank, J.: Modeling coupled unsaturated and saturated nitrate distribution of the aquifer Westliches Leibnitzer Feld, Austria, Environmental Earth Sciences, 69(2), 663–678, https://doi.org/10.1007/s12665-013-2302-6, 2013.

Kraus, D., Weller, S., Klatt, S., Haas, E., Wassmann, R., Kiese, R., and Butterbach-Bahl, K.. A new LandscapeDNDC biogeochemical module to predict CH4 and N2O emissions from lowland rice and upland cropping systems, Plant and Soil, 386(1–2), 125–149, https://doi.org/10.1007/s11104-014-2255-x, 2014, 2014.

Kroes, J. G., and Dam, J. C. van.: Reference Manual SWAP; version 3.0.3, 2003.

Kronvang, B., Borgvang, S. A., and Barkved, L. J.: Towards European harmonised procedures for quantification of nutrient losses from diffuse sources - The EUROHARP project. In Journal of Environmental Monitoring (Vol. 11, Issue 3, pp. 503–505). Royal Society of Chemistry. https://doi.org/10.1039/b902869m, 2009.

Kuhn, T.: The revision of the German Fertiliser Ordinance in 2017 The revision of the German Fertiliser Ordinance in 2017

Till Kuhn, 2017.

Kumar, M., Ou, Y. L., and Lin, J. G.: Co-composting of green waste and food waste at low C/N ratio. Waste Management, 30(4), 602–609. https://doi.org/10.1016/j.wasman.2009.11.023, 2010.

Lamarque, P., Tappeiner, U., Turner, C., Steinbacher, M., Bardgett, R. D., Szukics, U., Schermer, M., and Lavorel, S.: Stakeholder perceptions of grassland ecosystem services in relation to knowledge on soil fertility and biodiversity, Regional Environmental Change, 11(4), 791–804. https://doi.org/10.1007/s10113-011-0214-0, 2011.

Li, C. S.: Modeling trace gas emissions from agricultural ecosystems. In Methane Emissions from Major Rice Ecosystems in Asia (pp. 259–276). Springer Netherlands. https://doi.org/10.1007/978-94-010-0898-3_20, 2000.

Li, C., Salas, W., Zhang, R., Krauter, C., Rotz, A., and Mitloehner, F.: Manure-DNDC: A biogeochemical process model for quantifying greenhouse gas and ammonia emissions from livestock manure systems, Nutrient Cycling in Agroecosystems, 93(2), 163–200, https://doi.org/10.1007/s10705-012-9507-z, 2012.

Liu, S. M., Xu, Z. W., Wang, W. Z., Jia, Z. Z., Zhu, M. J., Bai, J., and Wang, J. M.: A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy balance closure problem, Hydrology and Earth System Sciences, 15(4), 1291–1306, https://doi.org/10.5194/hess-15-1291-2011, 2011.

Lü, X. T., Dijkstra, F. A., Kong, D. L., Wang, Z. W., and Han, X. G.: Plant nitrogen uptake drives responses of productivity to nitrogen and water addition in a grassland, Scientific Reports, 4(1), 1–7. https://doi.org/10.1038/srep04817, 2014.

Mahowald, N., Jickells, T. D., Baker, A. R., Artaxo, P., Benitez-Nelson, C. R., Bergametti, G., Bond, T. C., Chen, Y., Cohen, D. D., Herut, B., Kubilay, N., Losno, R., Luo, C., Maenhaut, W., McGee, K. A., Okin, G. S., Siefert, R. L., and Tsukuda, S.: Global distribution of atmospheric phosphorus sources, concentrations and deposition rates, and anthropogenic impacts. Global Biogeochemical Cycles, 22(4), https://doi.org/10.1029/2008GB003240, 2008

Manoli, G., Ivanov, V. Y., and Fatichi, S.: Dry-Season Greening and Water Stress in Amazonia: The Role of Modeling Leaf Phenology. Journal of Geophysical Research: Biogeosciences, 123(6), 1909–1926. https://doi.org/10.1029/2017JG004282, 2018.

Manzoni, S., Moyano, F., Kätterer, T., and Schimel, J.: Modeling coupled enzymatic and solute transport controls on decomposition in drying soils, Soil Biology and Biochemistry, 95, 275–287. https://doi.org/10.1016/j.soilbio.2016.01.006, 2016.

Marcolla, B., Cescatti, A., Manca, G., Zorer, R., Cavagna, M., Fiora, A., Gianelle, D., Rodeghiero, M., Sottocornola, M., and Zampedri, R.: Climatic controls and ecosystem responses drive the inter-annual variability of the net ecosystem exchange of an alpine meadow, Agricultural and Forest Meteorology, 151(9), 1233–1243. https://doi.org/10.1016/j.agrformet.2011.04.015, 2011.

Mastrotheodoros, T., Pappas, C., Molnar, P., Burlando, P., Hadjidoukas, P., and Fatichi, S.: Ecohydrological dynamics in the Alps: Insights from a modelling analysis of the spatial variability, Ecohydrology, 12(1). https://doi.org/10.1002/eco.2054, 2019.

Mastrotheodoros, T., Pappas, C., Molnar, P., Burlando, P., Manoli, G., Parajka, J., Rigon, R., Szeles, B., Bottazzi, M.,

675     Hadjidoukas, P., and Fatichi, S.: More green and less blue water in the Alps during warmer summers. Nature Climate
        Change, 10(2), 155–161, https://doi.org/10.1038/s41558-019-0676-5, 2020.

        Mauder, M., Foken, T., and Cuxart, J.: Surface-Energy-Balance Closure over Land: A Review. Boundary-Layer Meteorolgy,
        https://doi.org/10.1007/s10546-020-00529-6, 2020.

        Mauder, M., Genzel, S., Fu, J., Kiese, R., Soltani, M., Steinbrecher, R., Zeeman, M., Banerjee, T., De Roo, F., and Kunstmann,
680     H.: Evaluation of energy balance closure adjustment methods by independent evapotranspiration estimates from
        lysimeters and hydrological simulations. Hydrological Processes, 32(1), 39–50. https://doi.org/10.1002/hyp.11397,
        2018.

        Mauder, M., Liebethal, C., Göckede, M., Leps, J. P., Beyrich, F., and Foken, T.: Processing and quality control of flux data
        during LITFASS-2003. *Boundary-Layer Meteorology*, *121*(1), 67–88. https://doi.org/10.1007/s10546-006-9094-0,
685     2006.

        Merbold, L., Eugster, W., Stieger, J., Zahniser, M., Nelson, D., and Buchmann, N.: Greenhouse gas budget ($CO_2$, $CH_4$ and
        $N_2O$) of intensively managed grassland following restoration, Global Change Biology, 20(6), 1913–1928,
        https://doi.org/10.1111/gcb.12518, 2014.

        Migliavacca, M., Galvagno, M., Cremonese, E., Rossini, M., Meroni, M., Sonnentag, O., Cogliati, S., Manca, G., Diotri, F.,
690     Busetto, L., Cescatti, A., Colombo, R., Fava, F., Morra di Cella, U., Emiliano, P., Consolata, S., and Richardson, A. D.:
        Using digital repeat photography and eddy covariance data to model grasslandphenology and photosynthetic CO2
        uptake, Agricultural and Forest Meteorology, 151, 1325–1337, 2011.

        Millar, D. J., Ewers, B. E., Mackay, D. S., Peckham, S., Reed, D. E., and Sekoni, A.: Improving ecosystem-scale modeling of
        evapotranspiration using ecological mechanisms that account for compensatory responses following disturbance. Water
695     Resources Research, 53(9), 7853–7868. https://doi.org/10.1002/2017WR020823, 2017.

        Moorhead, D. L., Sinsabaugh, R. L., Linkins, A. E., and Reynolds, J. F.: Decomposition processes: Modelling approaches and
        applications. *Science of the Total Environment*, *183*(1–2), 137–149. https://doi.org/10.1016/0048-9697(95)04974-6,
        1996.

        Niklaus, P. A., Wardle, D. A., and Tate, K. R.: Effects of plant species diversity and composition on nitrogen cycling and the
700     trace gas balance of soils. Plant and Soil, 282(1–2), 83–98. https://doi.org/10.1007/s11104-005-5230-8, 2006.

        Nyamangara, J., Piha, M. I., and Kirchmann, H.: Interactions of aerobically decomposed cattle manure and nitrogen fertilizer
        applied to soil, In Nutrient Cycling in Agroecosystems (Vol. 54), 1999.

        Oberholzer, S., Prasuhn, V., and Hund, A.: Crop water use under Swiss pedoclimatic conditions – Evaluation of lysimeter data
        covering a seven-year period, Field Crops Research, 211, 48–65, https://doi.org/10.1016/j.fcr.2017.06.003, 2017.

705     Parton, W. J., Hartman, M., Ojima, D., and Schimel, D.: DAYCENT and its land surface submodel: Description and testing.
        Global and Planetary Change, 19(1–4), 35–48. https://doi.org/10.1016/S0921-8181(98)00040-X, 1998.

        Parton, W. J., Ojima, D. S., Cole, C. V., and Schimel, D. S.: A General Model for Soil Organic Matter Dynamics: Sensitivity
        to       Litter       Chemistry,       Texture       and       Management       (pp.       147–167),       John       Wiley       &       Sons,       Ltd.

https://doi.org/10.2136/sssaspecpub39.c9, 2015.

710  Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A.,
Humphrey, M., Isaac, P., Polidori, D., Ribeca, A., van Ingen, C., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö,
J., … Papale, D.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific
Data*, *7*(1), 225. https://doi.org/10.1038/s41597-020-0534-3, 2020.

Perego, A., Giussani, A., Sanna, M., and Fumagalli, M.: The ARMOSA simulation crop model: Overall features, calibration
715  and validation results Space-time mapping and modelling of soil properties in Mediterranean and Temperate areas View
project, In Article in Italian Journal of Agrometeorology, https://www.researchgate.net/publication/259615668, 2013.

Peukert, S., Griffith, B. A., Murray, P. J., Macleod, C. J. A., and Brazier, R. E.: Intensive Management in Grasslands Causes
Diffuse Water Pollution at the Farm Scale. Journal of Environmental Quality, 43(6), 2009–2023.
https://doi.org/10.2134/jeq2014.04.0193, 2014.

720  Phogat, V., Skewes, M. A., Cox, J. W., Alam, J., Grigson, G., and Šimůnek, J.: Evaluation of water movement and nitrate
dynamics in a lysimeter planted with an orange tree, Agricultural Water Management, 127, 74–84.
https://doi.org/10.1016/j.agwat.2013.05.017, 2013.

Prechsl, U. E., Burri, S., Gilgen, A. K., Kahmen, A., and Buchmann, N.: No shift to a deeper water uptake depth in response
to summer drought of two lowland and sub-alpine C3-grasslands in Switzerland, Oecologia, 177(1), 97–111.
725  https://doi.org/10.1007/s00442-014-3092-6, 2015.

Pütz, T., Kiese, R., Wollschläger, U., Groh, J., Rupp, H., Zacharias, S., Priesack, E., Gerke, H. H., Gasche, R., Bens, O., Borg,
E., Baessler, C., Kaiser, K., Herbrich, M., Munch, J., Sommer, M., Vogel, H., Vanderborght, J., Vereecken, H., … Pütz,
T.: TERENO-SOILCan: a lysimeter-network in Germany observing soil processes and plant diversity influenced by
climate change. Environmental Earth Sciences, *75*, 1242. https://doi.org/10.1007/s12665-016-6031-5, 2016.

730  Pütz, T., Fank, J., and Flury, M.: Lysimeters in Vadose Zone Research. Vadose Zone Journal, 17(1), 0.
https://doi.org/10.2136/vzj2018.02.0035, 2018.

Richter, G. M., Acutis, M., Trevisiol, P., Latiri, K., and Confalonieri, R.: Sensitivity analysis for a complex crop model applied
to Durum wheat in the Mediterranean. European Journal of Agronomy, 32(2), 127–136.
https://doi.org/10.1016/j.eja.2009.09.002, 2010.

735  Robertson, A. D., Paustian, K., Ogle, S., Wallenstein, M. D., Lugato, E., and Cotrufo, M. F.: Unifying soil organic matter
formation and persistence frameworks: the MEMS model, Biogeosciences, 16(6), 1225–1248.
https://doi.org/10.5194/bg-16-1225-2019, 2019.

Sala, O. E., and Paruelo, J. M.: Ecosystem services in grasslands. In Daily GC (Ed.), Nature's services: societal dependence
on natural ecosystems (pp. 237–251), 1997.

740  Saxton, K.E. and Rawls, W.J.: Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions,
Soil Sci. Soc. Am. J., 70: 1569-1578. doi:10.2136/sssaj2005.0117, 2006.

Schirpke, U., Kohler, M., Leitinger, G., Fontana, V., Tasser, E., and Tappeiner, U.: Future impacts of changing land-use and climate on ecosystem services of mountain grassland and their resilience, Ecosystem Services, 26, 79–94.

745     https://doi.org/10.1016/j.ecoser.2017.06.008, 2017.

Schoen, R., Gaudet, J. P., and Bariac, T.: Preferential flow and solute transport in a large lysimeter, under controlled boundary conditions. Journal of Hydrology, 215(1–4), 70–81. https://doi.org/10.1016/S0022-1694(98)00262-5, 1999.

Shajari, F., Einsiedl, F., and Rein, A.: Characterizing Water Flow in Vegetated Lysimeters with Stable Water Isotopes and Modeling, Groundwater, gwat.12970. https://doi.org/10.1111/gwat.12970, 2019.

750     Shi, Y., Davis, K. J., Duffy, C. J., and Yu, X.: Development of a Coupled Land Surface Hydrologic Model and Evaluation at a Critical Zone Observatory, Journal of Hydrometeorology, 14(5), 1401–1420, https://doi.org/10.1175/JHM-D-12-0145.1, 2013.

Siderius, C., Groenendijk, P., Gerven, L. P. A. Van, Jeuken, M. H. J. ., and Smit, A. A. M. F. R.: Process description of NuswaLite. A simplified model for the fate of nutrients in surface water. January, 71, 2009.

755     Simmelsgaard, S. E., and Djurhuus, J.: An empirical model for estimating nitrate leaching as affected by crop type and the long-term N fertilizer rate. *Soil Use and Management*, *14*(1), 37–43. https://doi.org/10.1111/j.1475-2743.1998.tb00608.x, 1998.

Smit, A. A. M. F. R., Siderius, C., and van Gerwen, L. P. A.: Process Description of SWQN. January, 2009.

Smith, W., Grant, B., Qi, Z., He, W., VanderZaag, A., Drury, C. F., and Helmers, M.: Development of the DNDC model to

760     improve soil hydrology and incorporate mechanistic tile drainage: A comparative analysis with RZWQM2. Environmental Modelling and Software, 123, 104577, https://doi.org/10.1016/j.envsoft.2019.104577, 2020.

Sohier, C., Degre, A., and Dautrebande, S.: From root zone modelling to regional forecasting of nitrate concentration in recharge flows–The case of the Walloon Region (Belgium). Elsevier. https://www.sciencedirect.com/science/article/pii/S0022169409001218, 2009.

765     Sommerfeldt, T. G., Chang, C., and Entz, T.: Long-term Annual Manure Applications Increase Soil Organic Matter and Nitrogen, and Decrease Carbon to Nitrogen Ratio, Soil Science Society of America Journal, 52(6), 1668–1672, https://doi.org/10.2136/sssaj1988.03615995005200060030x, 1988.

Spehn, E. M., Hector, A., Joshi, J., Scherer-Lorenzen, M., Schmid, B., Bazeley-White, E., Beierkuhnlein, C., Caldeira, M. C., Diemer, M., Dimitrakopoulos, P. G., Finn, J. A., Freitas, H., Giller, P. S., Good, J., Harris, R., Högberg, P., Huss-Danell,

770     K., Jumpponen, A., Koricheva, J., …, and Lawton, J. H.: Ecosystem effects of biodiversity manipulations in european grasslands, Ecological Monographs, 75(1), 37–63. https://doi.org/10.1890/03-4101, 2005.

Stenitzer, E.: SIMWASER. Ein numerisches Modell zur Simulation des Bodenwasserhaushaltes und des Pflanzenertrages eines Standortes. Petzenkirchen, 1988.

Swiss Federal Council: Verordnung vom 23. Oktober 2013 über die Direktzahlungen an die Landwirtschaft

775     (Direktzahlungsverordnung, DZV). https://www.admin.ch/opc/de/classified-compilation/20130216/index.html, 1998.

Tafteh, A., and Sepaskhah, A. R.: Application of HYDRUS-1D model for simulating water and nitrate leaching from

continuous and alternate furrow irrigated rapeseed and maize fields, Agricultural Water Management, 113, 19–29. https://doi.org/10.1016/j.agwat.2012.06.011, 2012.

Tague, C. L., McDowell, N. G., and Allen, C. D.: An Integrated Model of Environmental Effects on Growth, Carbohydrate Balance, and Mortality of Pinus ponderosa Forests in the Southern Rocky Mountains. *PLoS ONE*, *8*(11), e80286. https://doi.org/10.1371/journal.pone.0080286, 2013.

Tilman, D., Wedin, D., and Knops, J.: Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, *379*(6567), 718–720. https://doi.org/10.1038/379718a0, 1996.

Van Dam, J. C.: Field-scale water flow and solute transport: SWAP model concepts, parameter estimation and case studies. Wageningen University, 2000.

Velthof, G. L., Lesschen, J. P., Schils, R. L. M., Smit, A., Elbersen, B. S., Hazeu, G. W., Mucher, C. A., and Oenema, O.: Grassland areas, production and use Lot 2. Methodological studies in the field of Agro-Environmental Indicators, 2014.

Vescovo, L., and Gianelle, D.: Using the MIR bands in vegetation indices for the estimation of grassland biophysical parameters from satellite remote sensing in the Alps region of Trentino (Italy), Advances in Space Research, 41(11), 1764–1772. https://doi.org/10.1016/j.asr.2007.07.043, 2008.

Vet, R., Artz, R. S., Carou, S., Shaw, M., Ro, C. U., Aas, W., Baker, A., Bowersox, V. C., Dentener, F., Galy-Lacaux, C., Hou, A., Pienaar, J. J., Gillett, R., Forti, M. C., Gromov, S., Hara, H., Khodzher, T., Mahowald, N. M., Nickovic, S., …and Reid, N. W.: A global assessment of precipitation chemistry and deposition of sulfur, nitrogen, sea salt, base cations, organic acids, acidity and pH, and phosphorus, Atmospheric Environment, 93, 3–100, https://doi.org/10.1016/j.atmosenv.2013.10.060, 2014.

Widmoser, P., and Wohlfahrt, G.: Attributing the energy imbalance by concurrent lysimeter and eddy covariance evapotranspiration measurements. Agricultural and Forest Meteorology, 263, 287–291. https://doi.org/10.1016/j.agrformet.2018.09.003, 2018.

Wieder, W. R., Bonan, G. B., and Allison, S. D.: Global soil carbon projections are improved by modelling microbial processes. Nature Climate Change, 3(10), 909–912. https://doi.org/10.1038/nclimate1951, 2013.

Williams, J., Jones, C., and Dyke, P. T.: A modeling approach to determining the relationship between erosion and soil productivity, doi:10.13031/2013.32748, 1984.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B. E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., … Verma, S.: Energy balance closure at FLUXNET sites. Agricultural and Forest Meteorology, 113(1–4), 223–243. https://doi.org/10.1016/S0168-1923(02)00109-0, 2002.

Wohlfahrt, G., Anderson-Dunn, M., Bahn, M., Balzarolo, M., Berninger, F., Campbell, C., Carrara, A., Cescatti, A., Christensen, T., Dore, S., Eugster, W., Friborg, T., Furger, M., Gianelle, D., Gimeno, C., Hargreaves, K., Hari, P., Haslwanter, A., Johansson, T., … Cernusca, A.: Biotic, abiotic, and management controls on the net ecosystem CO2 exchange of European mountain grassland ecosystems. Ecosystems, 11(8), 1338–1351. https://doi.org/10.1007/s10021-

008-9196-2, 2008a.

Wohlfahrt, G., Hammerle, A., Haslwanter, A., Bahn, M., Tappeiner, U., and Cernusca, A.: Seasonal and inter-annual variability of the net ecosystem $CO_2$ exchange of a temperate mountain grassland: Effects of weather and management, Journal of Geophysical Research, 113(D8), https://doi.org/10.1029/2007JD009286, 2008b.

815 Wohlfahrt, G., Irschick, C., Thalinger, B., Hörtnagl, L., Obojes, N., and Hammerle, A.: Insights from Independent Evapotranspiration Estimates for Closing the Energy Balance: A Grassland Case Study. Vadose Zone Journal, 9(4), 1025–1033. https://doi.org/10.2136/vzj2009.0158, 2010.

Wolf, B., Chwala, C., Fersch, B., Garvelmann, J., Junkermann, W., Zeeman, M. J., Angerer, A., Adler, B., Beck, C., Brosy, C., Brugggger, P., Emeis, S., Dannenmann, M., De Roo, F., Diaz-Pines, E., Haas, E., Hagen, M., Hajnsek, I., Jacobeit,

820 J., …, and Schmid, H. P.: The scalex campaign: Scale-crossing land surface and boundary layer processes in the TERENO-prealpine observatory, Bulletin of the American Meteorological Society, 98(6), 1217–1234, https://doi.org/10.1175/BAMS-D-15-00277.1, 2017.

Yu, L., Zeng, S., Fatichi, S., and Su, Z.: How vadose zone mass and energy transfer physics affects the ecohydrological dynamics os a Tibetan meadow? *The Cryosphere Discussion*. https://doi.org/10.5194/tc-2020-88. 2020.

825 Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., …, and Vereecken, H.: A Network of Terrestrial Environmental Observatories in Germany, Vadose Zone Journal, 10(3), 955–973, https://doi.org/10.2136/vzj2010.0139, 2011.

Zeeman, M. J., Mauder, M., Steinbrecher, R., Heidbach, K., Eckart, E., and Schmid, H. P.: Reduced snow cover affects

830 productivity of upland temperate grasslands, Agricultural and Forest Meteorology, 232, 514–526, https://doi.org/10.1016/j.agrformet.2016.09.002, 2017.

Zeeman, M. J., Hiller, R., Gilgen, A. K., Michna, P., Plüss, P., Buchmann, N., and Eugster, W.: Management and climate impacts on net CO2 fluxes and carbon budgets of three grasslands along an elevational gradient in Switzerland, Agricultural and Forest Meteorology, 150(4), 519–530. https://doi.org/10.1016/j.agrformet.2010.01.011, 2010.

835 Zeeman, M. J., Shupe, H., Baessler, C., and Ruehr, N. K.: Productivity and vegetation structure of three differently managed temperate grasslands. Agriculture, Ecosystems and Environment, 270–271, 129–148. https://doi.org/10.1016/j.agee.2018.10.003, 2019.

Zhu, N.: Effect of low initial C/N ratio on aerobic composting of swine manure with rice straw, Bioresource Technology, 98(1), 9–13. https://doi.org/10.1016/j.biortech.2005.12.003, 2007.

840
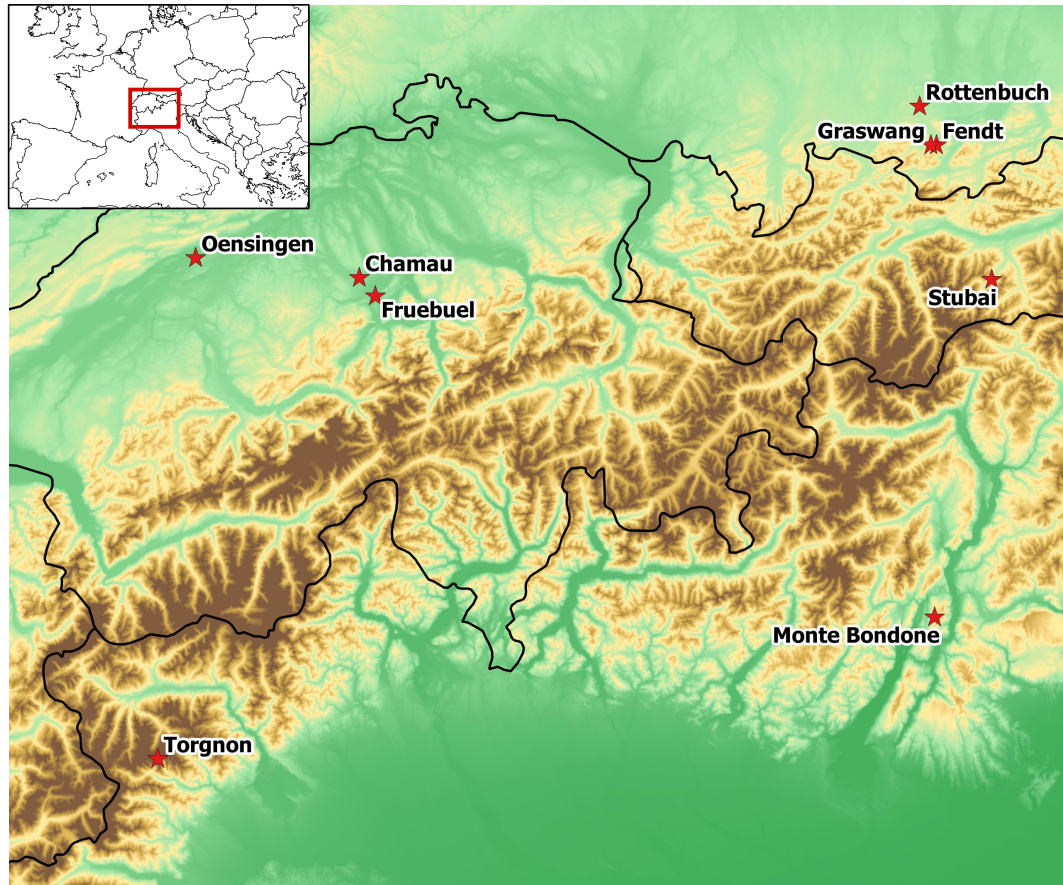
**Figure 1. Location of the study sites.** The 9 sites are situated across the European Alps in Italy (TOR, MB), Switzerland (CHA, OEN, FRU), Austria (STU) and Germany (FEN, ROT, GRA). The map was produced by the Authors and the map with countries' borders was retrieved from "Made with Natural Earth" and the DTM from SwissTopo.ch.
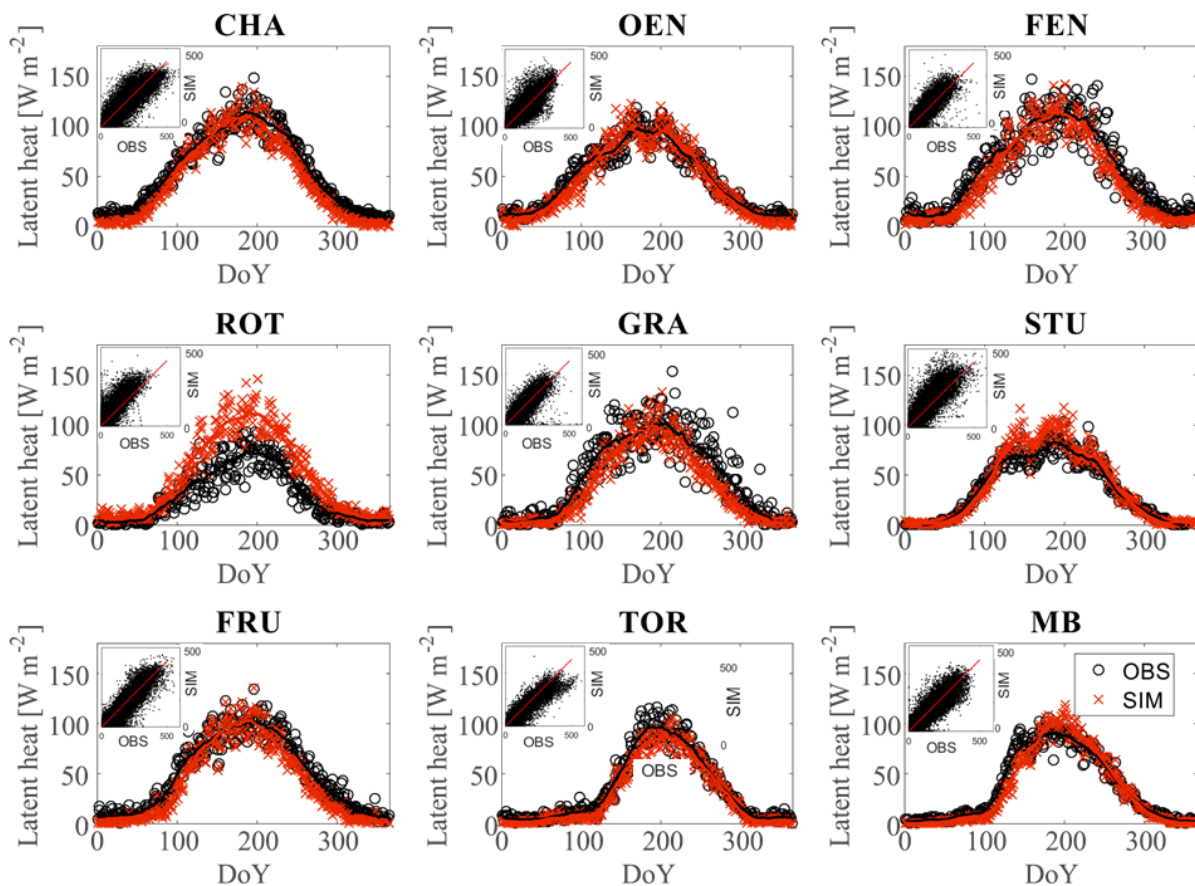
**Figure 2. Observed versus simulated seasonal daily latent heat fluxes.** We compare the observed (black points) and simulated (red crosses) seasonal pattern of latent heat computing the average value for every day of the year (DoY) considering all the years for which observations are available. We also apply a moving average with a centered window of 30 days (continuous lines). In the upper left corner of each subplot a scatter plot comparison of the hourly values of observed and simulated latent heat is shown. In red is plotted the 1:1 line.
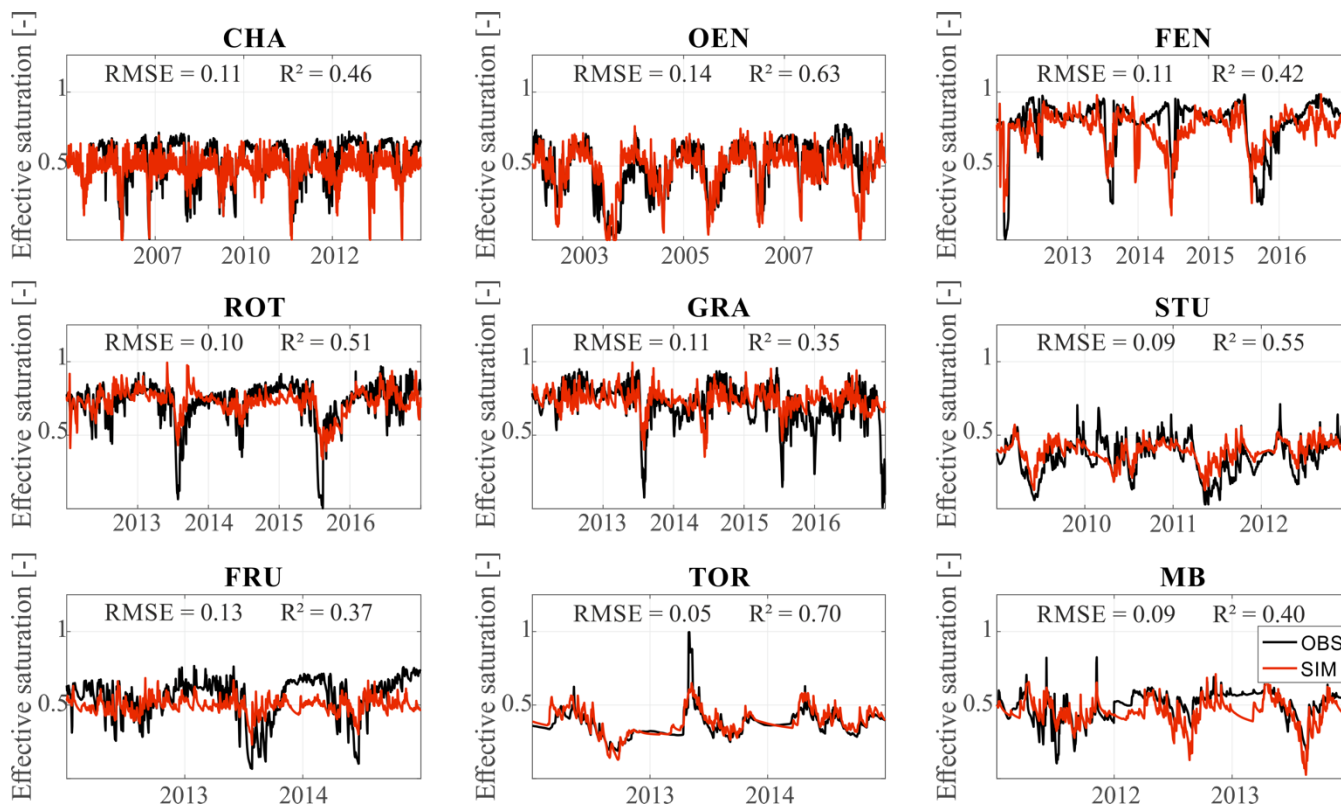
**Figure 3. Observed vs simulated daily effective saturation.** We compare the observed (black) and measured (red) pattern
855    of the effective saturation across all the sites. In CHA, STU, FRU, TOR soil water content is measured at 5 cm depth, in OEN
and MB at 10 cm depth and in FEN, ROT and GRA at 12 cm depth. The goodness of fit metrics Room Mean Square Error
(RMSE) and coefficient of determination ($R^2$) are reported for each station in each subplot.
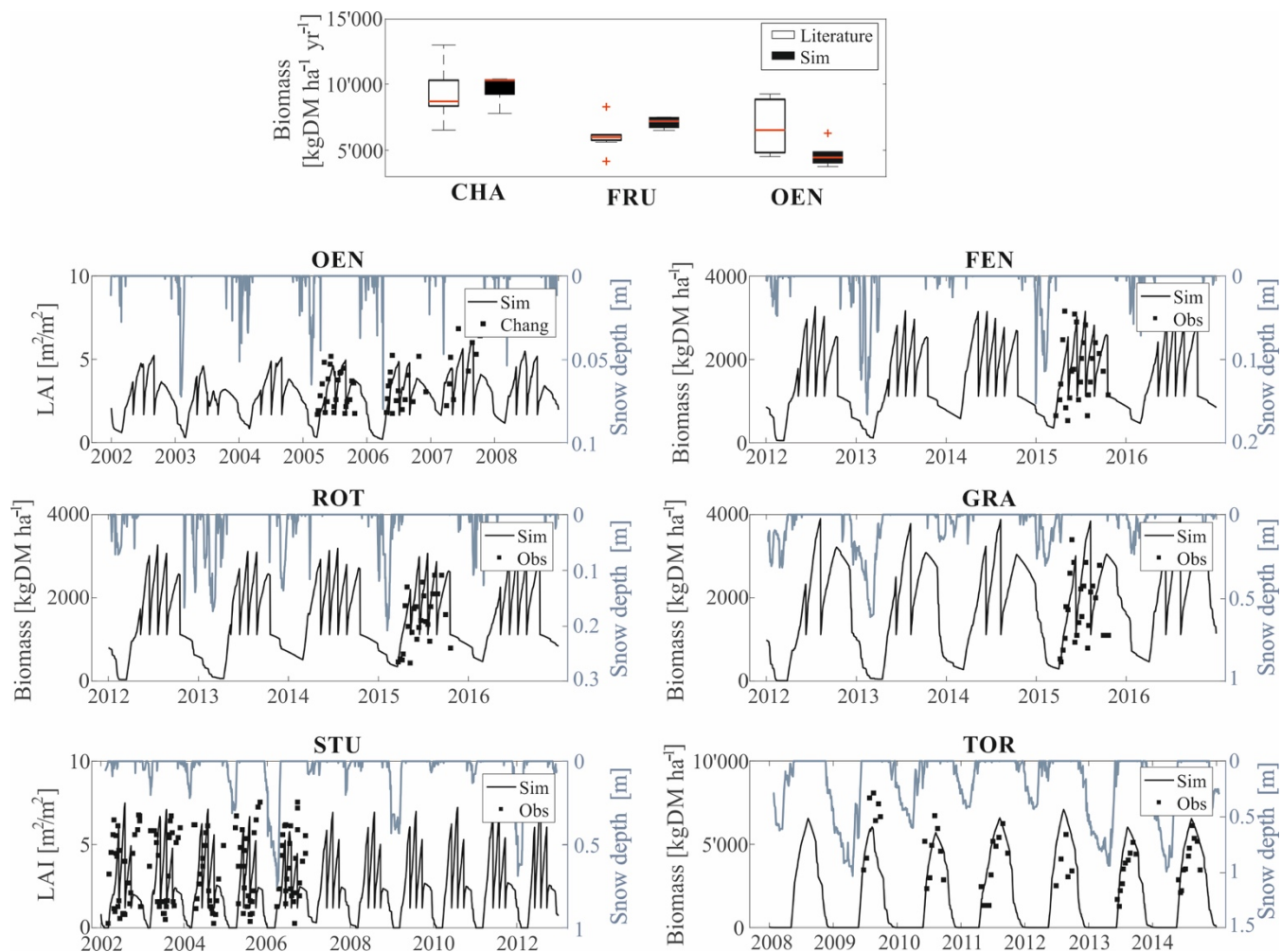
**Figure 4. Observed versus simulated leaf-biomass and LAI.** (a) The simulated mean yearly harvested biomass in CHA,
FRU and OEN are compared with published values. For CHA data are extracted from Gilgen and Buchmann (2009), Zeeman
et al. (2010) and Prechsl et al. (2015). For FRU we compare with data from Gilgen and Buchmann (2009) and Zeeman et al.
(2010). For the site OEN we compare simulated values with data reported by Ammann et al. (2009). (b) LAI in OEN is
compared with observations from Chang et al. (2013). (c)(d)(e) Biomass data for the sites in Germany FEN, ROT and GRA
were provided by the ScaleX campaign 2015 (Zeeman et al., 2019). (f) LAI data in STU were digitalized from Wohlfahrt et
al., 2008b (g) biomass data in TOR, observations were provided by the Environmental Protection Agency of Aosta Valley
(Filippa et al., 2015). In all the subplots, we compare the simulated either biomass or LAI (black line) with observations (black
dots), simulated snow depth (grey line) is also shown.
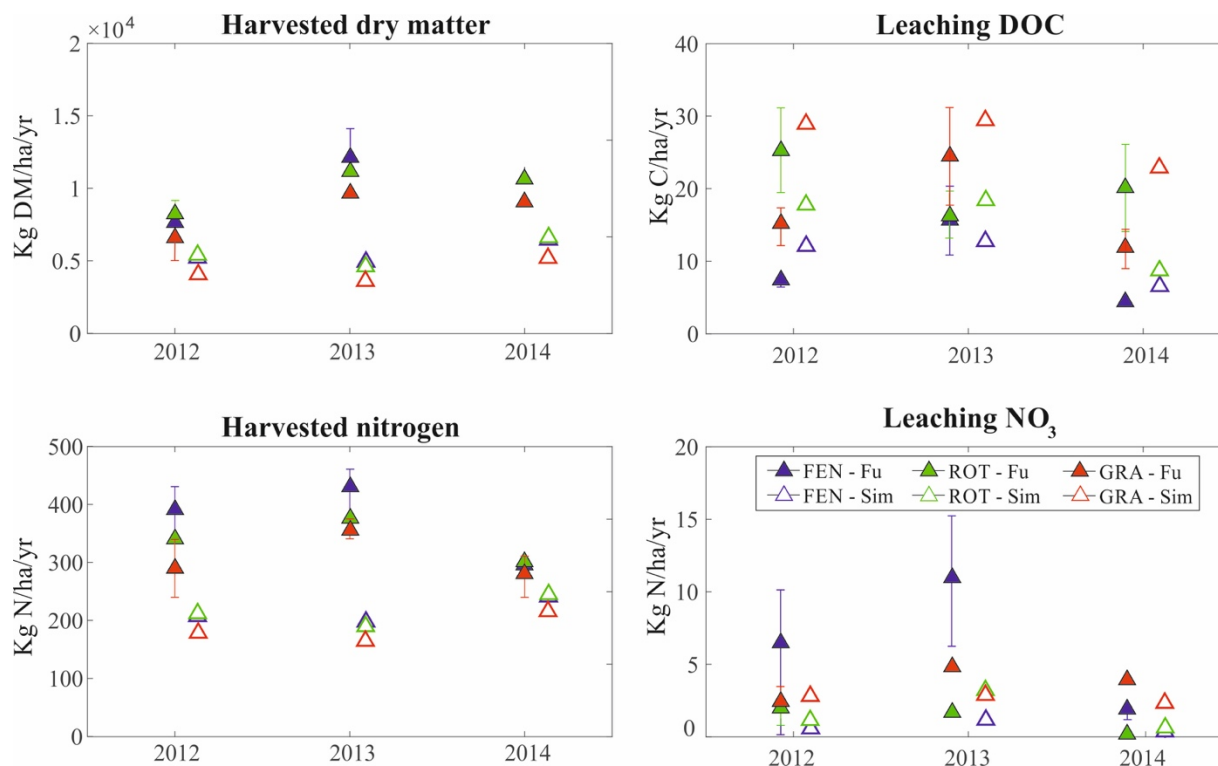
**Figure 5. Observed versus simulated harvested dry matter, harvested nitrogen, leaching of DOC and NO₃.** Simulated
870    annual totals (empty triangles) are compared with annual totals (full triangles) reported by Fu et al. (2017; 2019) in FEN (blue),
ROT (green) and GRA (red). The uncertainty bars represent the 25th and 75th percentile of observations.
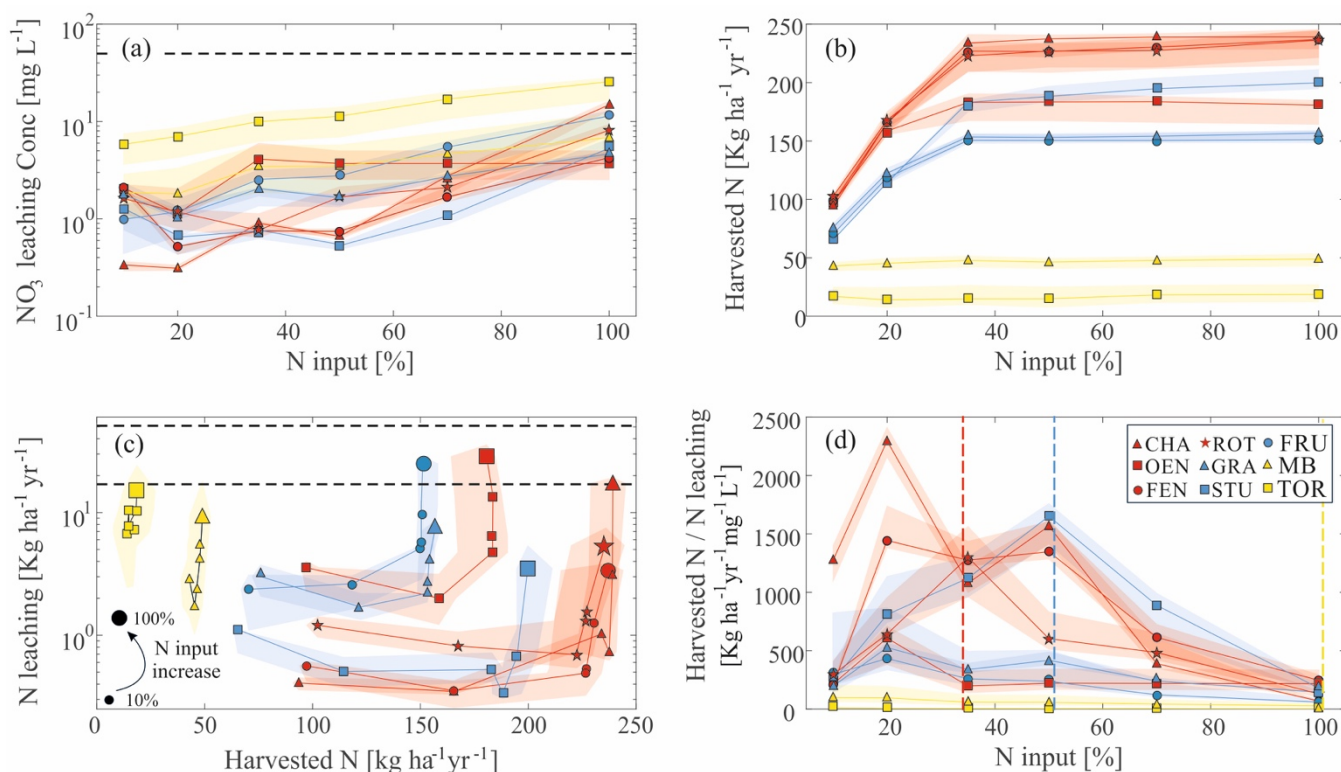
**Figure 6. Results of the fertilization experiments.** (a) NO₃ concentration in groundwater recharge as a function of different nitrogen fertilization scenarios (percentage of the maximum) in each site. (b) Harvested N in each site as function of different nitrogen fertilization scenarios. (c) Harvested N vs N leaching in each site. Moving counterclockwise follows the increase in N input. As a reference the biggest marker indicates N input of 100%. The dotted lines represent the estimate of maximum NO₃ losses assumed by EU regulations or nearby countries. They correspond to the values 17 and 51 kgN ha$^{-1}$ yr$^{-1}$, i.e., 10% and 30% of the maximum allowed input 170 kgN ha$^{-1}$ yr$^{-1}$. (d) N fertilization efficiency index computed as the ratio between the harvested N and N concentration in groundwater recharge as a function of nitrogen input. In all the subplots the colors represent the elevation class, i.e., pre-Alpine (red), Alpine (blue) and high-Alpine (yellow) sites. The colored area around the markers and lines represent the 25$^{th}$ and 75$^{th}$ percentile of the internannual variability of the simulated variables. The vertical dashed bars represent the limit of 170 kgN ha$^{-1}$ yr$^{-1}$ imposed by the EU Nitrate Directive in each of the three classes.
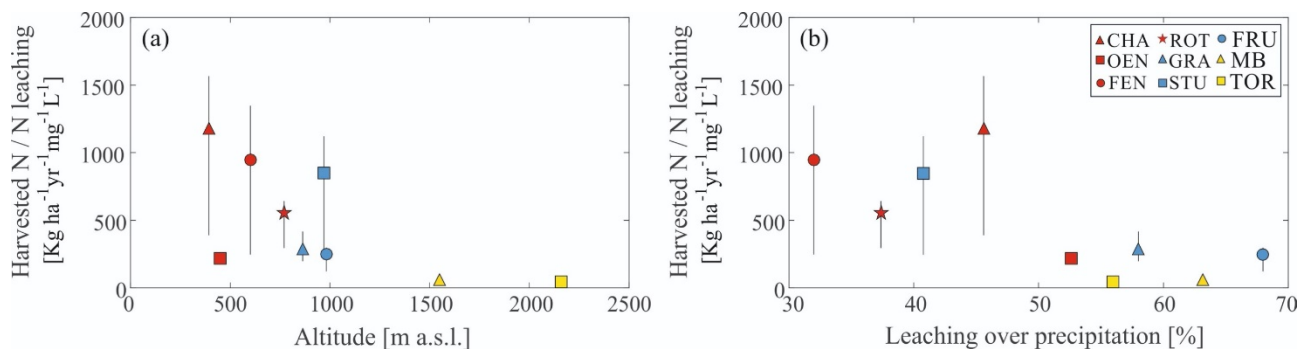
**Figure 7. Ratio between harvested N and N leaching concentration as a function of site characteristics.** (A) Harvested
885 N/ N leaching concentration as a function of elevation in each site. (B) Harvested N/ N leaching concentration as a function of
the percentage of yearly water recharge to groundwater over the yearly precipitation. In both plots the whiskers span the 25th
and 75th percentile of the interannual variability.

**Table 1. Site characteristics.** The following main sites characteristics are reported: elevation, latitude, longitude, period for
890 which measurements are available, Mean Annual Precipitation (MAP), mean air temperature (mean Ta), number of grass cuts
per year and main references to original studies referring to the study sites. The values of MAP and mean Ta are computed
from the available time series.

| Site | Elevation (m a.s.l.) | Latitude | Longitude | Data availability | MAP (mm) | Mean Ta (°C) | # cuts year⁻¹ | References |
|------|---------------------|----------|-----------|-------------------|----------|--------------|---------------|------------|
| CHA | 393 | 47°12' | 8°24' | 2006 - 2014 | 1'134 | 9.4 | 6 | Gilgen and Buchmann, 2009; Zeeman et al., 2010 |
| OEN | 450 | 47°17' | 7°44' | 2002 - 2008 | 1'222 | 9.2 | 3 | Ammann et al., 2007; 2009 |
| FEN | 600 | 47°57' | 11°06' | 2012 - 2016 | 850 | 7.8 | 5 | Kiese et al., 2018 |
| ROT | 770 | 47°70' | 10°98' | 2012 - 2016 | 955 | 7.4 | 5 | Kiese et al., 2018 |
| GRA | 864 | 47°57' | 11°03' | 2012 - 2016 | 1'114 | 5.3 | 2 | Kiese et al., 2018 |
| STU | 970 | 47°12' | 11°32' | 2002 - 2012 | 856 | 6.8 | 3 | Hammerle et al., 2008; Wohlfahrt et al., 2008a, 2010 |
| FRU | 982 | 47°60' | 8°32' | 2005 - 2014 | 1'627 | 7.6 | 4 | Gilgen and Buchmann, 2009; Zeeman et al., 2010 |
| MB | 1550 | 46°00' | 11°02' | 2003 - 2013 | 1'268 | 5.2 | 1 | Gilmanov et al., 2007; Vescovo and |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Gianelle, 2008; Gianelle et al., 2009; Marcolla et al., 2011 |
| **TOR** | 2160 | 45°50' | 7°34' | 2008 - 2015 | 870 | 2.9 | 0 | Migliavacca et al., 2011; Galvagno et al., 2013 Filippa et al., 2015 |

**Table 2. Numerical experiments.** The sites are divided into pre-Alpine, Alpine, and high-Alpine sites based on the elevation in which they are situated. For each class the number of cuts and manure applications as well as the yearly N-load injected in the different modelling experiments is reported. The specific manure application is the same for each site class, but the total amount varies across classes as the number of yearly applications is different. The results are presented as a function of the percentage of max N input, reported in the table for each experimental setup.

| Pre-Alpine/Intensive | Alpine/Extensive | High-Alpine/Extensive | | |
|---|---|---|---|---|
| CHA, OEN, FEN, ROT | GRA, STU, FRU | MB, TOR | | |
| 5 cuts, 6 fertilizations | 3 cuts, 4 fertilizations | 1 cut, 2 fertilizations | | |
| Yearly N load (kg ha$^{-1}$ yr$^{-1}$) | Yearly N load (kg ha$^{-1}$ yr$^{-1}$) | Yearly N load (kg ha$^{-1}$ yr$^{-1}$) | Specific manure application (kg ha$^{-1}$ yr$^{-1}$) | % of max N input |
| 50 | 33 | 17 | 74 | 10% |
| 100 | 67 | 33 | 148 | 20% |
| 175 | 117 | 58 | 260 | 35% |
| 250 | 167 | 83 | 371 | 50% |
| 350 | 233 | 117 | 519 | 70% |
| 500 | 333 | 167 | 742 | 100% |

**Table 3. Coefficient of determination ($R^2$) of the simulated vs observed energy and carbon fluxes from flux towers data.**

| | Net radiation | Sensible heat | Latent heat | GPP |
|---|---|---|---|---|
| CHA | 0.98 | 0.63 | 0.86 | 0.78 |
| OEN | 0.98 | 0.66 | 0.82 | 0.69 |
| FEN | 0.97 | 0.51 | 0.79 | 0.66 |
| ROT | 0.74 | 0.48 | 0.76 | 0.63 |

| | | | |
|---|---|---|---|
| GRA | 0.74 | 0.39 | 0.70 | 0.59 |
| STU | 0.95 | 0.53 | 0.85 | 0.83 |
| FRU | 0.96 | 0.72 | 0.87 | 0.85 |
| MB | 0.93 | 0.75 | 0.88 | 0.79 |
| TOR | 0.94 | 0.69 | 0.87 | 0.72 |

900

**Table 4. Simulated energy, water and carbon fluxes at each site.** The mean annual evapotranspiration (ET), mean annual net radiation (Rn), mean annual Bowen ratio (Bo), mean annual gross primary production (GPP), mean gross primary production of July (July GPP) and the mean day of the year in which the growing season starts (Start growing season) are reported.

| | ET (mm yr$^{-1}$) | Rn (W m$^{-2}$) | Bo (/) | GPP (gC m$^{-2}$ yr$^{-1}$) | July GPP (gC m$^{-2}$ month$^{-1}$) | Start growing season (DoY) |
|---|---|---|---|---|---|---|
| CHA | 606 | 67.7 | 0.35 | 2177 | 362 | 76 |
| OEN | 581 | 74.8 | 0.47 | 1856 | 271 | 75 |
| FEN | 621 | 69.1 | 0.30 | 1752 | 309 | 88 |
| ROT | 617 | 72.1 | 0.26 | 1702 | 310 | 97 |
| GRA | 497 | 60.4 | 0.40 | 1422 | 284 | 113 |
| STU | 503 | 53.1 | 0.33 | 1574 | 353 | 99 |
| FRU | 517 | 59.6 | 0.36 | 1701 | 361 | 106 |
| MB | 433 | 61.9 | 0.62 | 1326 | 222 | 174 |
| TOR | 394 | 68.6 | 0.72 | 946 | 317 | 160 |

905