Response to RC1

We thank the anonymous reviewer for the kind words and positive assessment of the submitted manuscript. Below we first reply to the two major points raised and follow with a detailed line-by-line response to the minor comments (responses in blue).

Major comment 1: Manuscript section organization and improvement of validation

It is a bit strange to find the validation of the model near the end of the paper, in the Discussion section (Evaluation of predicted PNV), and furthermore this validation should be more quantitative. This section should come much earlier in the paper, probably at the beginning of the Results section. Other parts of the Discussion section could also be transferred to the Results section (but maybe near the end of the section): the sensitivity tests performed in subsections 5.3 and 5.4; these are results and not just discussion. More importantly, I feel that the validation of subsection 5.1 should be improved. As it is now, it is limited to a visual comparison of biome maps, as well of the foliage projection cover map predicted by the model with the vegetation cover map derived from MODIS data. You should provide at least some statistics for this comparison. Also, there is no validation of runoff, while it is reported as a very important variable for the landscape evolution model.

The organization of result and discussion section could have been clearer (as also noticed by Reviewer 2). To improve the readability we restructured parts of the result and discussion section. The evaluation of PNV distribution (biome locations) for PD conditions was merged into section 4.2. The comparison of simulated foliar projected cover with MODIS-based vegetation cover estimates was combined in result section 4.3. Furthermore, we introduced a new section that describes the results from our CO_2 sensitivity simulation experiment (formerly covered in 5.3). We did however keep the section describing the results from an alternative paleo climate dataset (ECHAM5) in the discussion section, as its not part of the TraCE-21ka-based simulation ensemble of this paper and was merely included to illustrate the importance of climate data for DVM simulation results.

Furthermore, we added more quantifications of discrepancies between modelled and observed FPC to the relevant section (mean absolute error, MAE). However, we want to note that the evaluation of both FPC and biome distribution is complicated for multiple reasons (which was also our reason for our simplistic visual characterization of the results). First, mismatches of biome classifications can be the result from multiple causes. The biome map given in Fig. 1a is based on an aggregated version of the floristic units (Luebbert and Pliscoff, 2017) that do not necessarily align with a PFT- and physiological-based classification. Also the large number units and their often very specific mosaic of co-occurring species make a clear separation into major biomes (that are based on very few characteristic PFTs) often impossible. For instance, Matorral and Sclerophyllous Woodland were not easily separable based on these units.

A thorough runoff validation is unfortunately not possible in the scope of this study (esp. due to a lack of spatially explicit runoff data). However, as Part II (Schmid et al., 2018) only uses FPC yet, we opted to postpone this issue to the planned future publication of the coupled DVM-LEM model which will include FPC and runoff effects. We thus added a paragraph that acknowledges this limitation and we revised all sections that concern runoff results. We still want to include runoff data is it will be considered in future work (as mentioned).

We added a quantitative assessment of simulated FPC against MODIS vegetation cover (Sect. 4.3) and correlation analysis when appropriate.

Major comment 2: Request for more details of landform effect in simulations

It is not clear to me that this study fulfils the objective of demonstrating the ability of a dynamic vegetation model to produce results useful for the spatial and temporal scales of landscapes evolution models. The authors develop a landform sub-model, but they do not really test it. They just present some transient evolution for one landform in each of four given model pixels. However, we do not know how far the use of the landform sub-model improves model prediction. It would be useful to illustrate the landform results for a given pixel (in addition to the altitudinal profiles that I guess use the landforms).
The reason to present only a single landform in Fig. 9 was to clearly show the PFT transitions with time. If one reports the average PFT composition of all landforms in a grid cell these trends are masked/ harder to showcase (but all simulations reported in the manuscript are run using the landform approach).
Another reason for originally not including the full details of the implemented landform approach was that we assumed this would bring the document to an unfeasible length. However, since both reviewers ask for more details of the approach we now include a new Appendix C that explains the conceptual approach and the modifications of site conditions for the landforms in detail. The landform simulation mode in general improved the simulation especially in semi-arid to Mediterranean locations as a heterogeneous landscape representation (i.e. valley landforms and higher altitude locations with lower temperatures) generally led to higher vegetation cover (access to more soil water, lower temperature with less water-limiting conditions).

We included a new figure to the supplement that highlights the differences in simulated FPC for individual landforms, the area-weighted average FPC from those landforms and the original LPJ-GUESS simulation for the last 2000 years of the transient simulation runs of the four focus sites.
As can be observed (Fig. S4a), the implemented landform approach has differing effects for the four focus sites. The area-weighted average FPC at site Sta. Gracia closely resembles the results from the default simulation mode (apart from landform 810 of high altitudes that only covers 1.7% of the grid cell area, Fig. S4b). Average-landform and default results for site La Campana also differ only marginally. However, here a set of landforms of higher altitudes has a substantially lower FPC than the average (~ - 15%). Variation at the hyper-arid site Pan de Azucar is lower (as is the FPC), but generally higher than the default simulation (which aligns with MODIS observations for the site, where the default model underestimates satellite-observed cover). The higher FPC in the new model setup is likely a result of deeper soil profiles of flat and valley landforms that allow a longer water storage versus the default uniform 1.5m soil assumption of LPJ-GUESS. The larger variability of FPC at site Nahuelbuta can be attributed to the rel. Large altitudinal variation in this 0.5x0.5 grid cell (coast to mountainous terrain) and is thus likely a temperature effect.
From site explorations (see Fig. 1 for impressions from the four locations) it is clear that vegetation is not distributed uniformly in the landscapes. Thus, the higher spatial diversity of simulated FPC in the landform approach can be assumed to more realistically describe true FPC in these areas and should thus also lead to more non-uniform erosion rates when FPC is spatially disaggregated on a high-resolution landscape utilized by an LEM.

We added a section (5.1) to the discussion section that points at these results.

Also the results of simulations with the landforms should be compared to those obtained with a model without landforms. How far does it improve the comparison with observed vegetation, or with MODIS vegetation cover? How far the results are affected by the adjustment of radiation for slope and aspects, or by the change in soil depth from the valley to the mountain

slope or ridge? Landform modelling is a novel aspect put forward by this paper. So, it is important to discuss it more fully.
We acknowledge that a thorough analysis of individual effects on simulated FPC would be interesting, but we deem this outside the scope of this manuscript as it would lead to a substantial extension of the paper which is already very long. However, we added section that discusses the observed differences to default model simulations (see answer above).

Minor comments

P. 2 lines 15-20: this paragraph provides a review of the use of DGVMs for paleoclimatic applications. They, however, mostly refer to studies performed with LPJ-GUESS. Please, please provide also examples of studies performed with other DGVMs.
LPJ-GUESS related paper did indeed dominate this section due to the widespread use of the model for these kind of simulations. We added the references Bragg et al., 2013 (BIOME4 model), Cowling et al., 2008 (LGM to PD simulation study for Africa using TRIFFID), Hopcroft et al., 2017 (a multi model study for the Holocene Sahara greening) to provide results from other models. In addition, we replace Shellito and Sloan (2006) Part 1 with the companion paper Part 2 since it investigates the possibilities DVM in more detail (the authors use the NCAR LSM-DGVM).
Furthermore, we added Snell et al. (2014) in the introduction paragraph to DGVMs as it is a nice review paper for readers new to the field (the authors also discuss multiple models and their strong points and weaknesses).

Bragg, F. J., Prentice, I. C., Harrison, S. P., Eglinton, G., Foster, P. N., Rommerskirchen, F., and Rullkötter, J.: Stable isotope and modelling evidence for CO_2 as a driver of glacial–interglacial vegetation shifts in southern Africa, Biogeosciences, 10, 2001-2010, https://doi.org/10.5194/bg-10-2001-2013, 2013

Cowling, S. A., Cox, P. M., Jones, C. D., Maslin, M. A., Peros, M., and Spall, S. A.: Simulated glacial and interglacial vegetation across Africa: implications for species phylogenies and trans-African migration of plants and animals. Global Change Biology, 14: 827-840, doi:10.1111/j.1365-2486.2007.01524.x, 2008.

Hopcroft, P. O., P. J. Valdes, A. B. Harper, and D. J. Beerling (2017), Multi vegetation model evaluation of the Green Sahara climate regime, *Geophys. Res. Lett.*, *44*, 6804–6813, doi:10.1002/2017GL073740.

Shellito, C. J. and Sloan, L. C.: Reconstructing a lost Eocene Paradise, Part II: On the utility of dynamic global vegetation models in pre-Quaternary climate studies, Glob. Planet. Change, 50(1), 18-32, doi: 10.1016/j.gloplacha.2005.08.002, 2006.

Snell, R. S., Huth, A. , Nabel, J. E., Bocedi, G. , Travis, J. M., Gravel, D. , Bugmann, H. , Gutiérrez, A. G., Hickler, T. , Higgins, S. I., Reineking, B. , Scherstjanoi, M. , Zurbriggen, N., and Lischke, H.: Using dynamic vegetation models to simulate plant range shifts. Ecography, 37: 1184-1197. doi:10.1111/ecog.00580, 2014.

P. 4, line 25: “We approximate the fraction A of the land surface . . .” instead of “We approximate the land surface ...”
Corrected

P. 4, line 37: Field capacity looks strange here. This would mean that a bucket approach is used in both layers. However, since drainage is not possible below field capacity (this is its definition), it would mean that subsurface runoff and percolation rate through the second layer are always zero in your model. Please check

LPJ-GUESS uses indeed a bucket model (Gerten et al., 2004; Smith et al., 2014; Seiler et al., 2015). The relevant section from Gerten et al., 2004 p254: "The model diagnoses surface runoff (R1) and subsurface runoff (R2) from the excess of water over field capacity of the upper and the lower soil layer, respectively. In addition, the amount of water percolating through the second soil layer is assumed to contribute to subsurface runoff (…)"

Baseflow is not explicitly mentioned in Gerten et al., 2004, however Seiler et al. 2015 state: "Precipitation enters the soil until the upper layer is saturated, while any additional precipitation is lost as surface runoff. Soil water evaporates from the upper 20cm, depending on potential evaporation and soil water content. Soil water percolates from the upper to the lower soil layer, until the lower soil layer is saturated, in which case excess water is lost as drainage. Water contained in the lower soil layer can leave the soil as baseflow at a given rate."

We rephrased the paragraph to: ". In LPJ-GUESS water enters the top soil layer as precipitation until this layer is fully saturated (excess water is lost as surface runoff and evaporation removes water from a 20cm sub-horizon of the top layer). During precipitation days, water can percolate from the top to the lower layer until the lower layer is saturated (excess water is lost as drainage). In addition, water of the lower layer can drain as baseflow with a fixed drainage rate (Gerten et al., 2004; Seiler et al., 2015). The model does neither consider lateral water movement between grid cells nor routing in a stream network (in this study we report the surface runoff component only)."

Seiler, C., R. W. A. Hutjes, B. Kruijt, and T. Hickler (2015), The sensitivity of wet and dry tropical forests to climate change in Bolivia. *J. Geophys. Res. Biogeosci.*, 120, 399–413. doi: 10.1002/2014JG002749.

P. 5, line 15: you use a constant average lapse rate of 6.5◦C/km, whereas the lapse rate could significantly vary, especially in desert areas where it could tend towards the dry adiabatic value of 9.7◦C/km. Moreover, other climate variables can change significantly with elevation in mountain areas, such as precipitation, cloudiness and air relative humidity.

We use 6.5 °C/km as it is an accepted global standard average value used by the climate science community. The reason is that this value is close to the global average, and is also the defined lapse rate in the International Standard Atmosphere (ISA) (e.g. Vaughan 2015). However, we acknowledge that the lapse rate varies a lot in space and time over multiple time scales (ranging from sub-daily to climatological). The high (spatial and temporal) variability is attributed to many features, associated with both atmospheric therodynamics and dynamics, e.g. radiative conditions, moisture content and large-scale atmospheric circulation. Hence, while a higher lapse would potentially be a better approximation for drier site (e.g. Pan de Azucar), this might not be the case for other sites with different atmospheric conditions. In addition, in our case the lapse rate correction is applied for the surface air temperature, which means that we would need to account for the surface conditions (e.g. vegetation type and potential snow cover) when estimating the lapse rate. To study the behavior of the near-surface lapse rate would require details of the atmospheric boundary layer on sub-daily to seasonal time scales. Although this would be an interesting exercise on its own, it is well outside the scope of this study. Another complicated issue is that we are dealing with paleo conditions. Unfortunately, there are no observational proxy records of past lapse rates. However, it is likely to believe that past lapse rates were different compared to the present because of differences in the atmospheric circulation, as well as radiative and surface conditions. While a value close to the dry adiabat (7-9 C/km) of the near-surface lapse rate might be a good approximation for present deserts, it is not certain that this is true also for past climates with different insolation conditions. For example, during episodes of lower insolation it is possible that the surface would be significantly cooler on average, and hence force the near-surface lapse rate toward lower values. Hence, since we cannot account for all possible uncertainties related to the spatial and temporal variations of the lapse rate, we decided to use one recognized value (the standard lapse rate) for all the sites in the study. However, we did include some of this discussion as well as potential implications of the chosen lapse rate in the revised manuscript.

Vaughan, W. W.: BASIC ATMOSPHERIC STRUCTURE AND CONCEPTS, Standard Atmosphere, 12-16, 2015.

P. 6, section 3.4: for PD, you use the 1960-1989 period. Does the atmospheric CO2 for PD correspond to the mean CO2 during this period? If so, it is significantly larger than the Holocene mean value and it is thus necessary to perform a pre-industrial simulation in addition to PD, in order to separate the CO2 and the climate effect in the difference between PD and MH.
We use annual CO_2 concentrations throughout the simulations (transient and time-slice simulations) and the preceding spin-up periods. Indeed, this concentration is higher than the Holocene average, but we fail to see how this makes a Preindustrial time-slice run necessary. We assess the impact of atmospheric CO_2 concentrations in the LGM evaluation runs (identical temperature/ precipitation regime; modified CO_2 concentration), but do not aim to separate temperature and CO2 effects on for PD and MH periods. We did however refer to the relevant paragraph in the results section and mention a possible effect of CO_2 in the observed differences between FPC for various time slices (MH – PD).

P. 7, line 20: “the Deciduous ‘Maule’ Forest occurs as total rainfall decreases and rainfall seasonality increases.” According to table 1, the ‘Maule’ Forest is a temperate forest made of temperate summergreen trees, not raingreen trees. So, we would expect that it is the seasonality of temperature and not rainfall that determines the occurrence of these trees. Please be more precise on the processes that link this forest to rainfall seasonality.
The reviewer is correct that the TeBS PFT variants that define the DMF biome are of summergreen phenology (species: *Nothofagus glauca*, *N. obliqua, N. alessandrii*). Especially *N. obliqua* is a very versatile species in the mesotemperate climate and covers large areas (Amigo & Rodríguez Guitián, 2011). Reviewer 2 is also correct that the sentence was not very clearly worded as we only mentioned that they occur north of the Valdivian evergreen forests (that receive higher rainfall that the areas where TeBS dominate). However, clearly, temperature and temp. seasonality is a driving factor for the emergence of these species at these latitudes. We thus rephrased the sentence to make this clearer.

P. 8 lines 19-21 and table 3: FPC in the south is lower during MH than at PD. Why? According to Fig. 3, in the south, the climate is wetter and colder during MH. We would expect larger FPC. Is the difference due to CO2, which is larger at PD? This needs to be commented.
According to the TraCE climate data, average MH temperature at latitudes 45-53 S was 0.5 deg C colder then PD. Precipitation totals were however higher from 35-46 S and lower for the southern areas (47-54). A reduced FPC for latitudes south of 45 (Fig. 6b) thus align more with the temperature difference. Furthermore, the PFT distribution maps of MH and (Fig. S1) and PD (Fig. 4) indicate that most reduction of FPC might be attributed to a smaller extent of temperate broadleaf evergreen PFTs (TeBE_tm, TeBE_itm) due to them being outcompeted on lower temperatures by the boreal PFT types. While lower CO_2 concentrations could also lead

to reduced FPC as illustrated by the LGM CO_2 sensitivity simulations (Fig. 10), we do not observe a general FPC reduction (however CO_2 concentration is lower for all areas at MH). Thus, differences in FPC are likely due to changes in PFT composition (FPC per PFT depends on the PFT properties, the balance between PFT compositions can vary with little changes in environmental envelope and are also a result of successional establishment) and/ or the differences in [CO_2] (as different PFTs can benefit from higher CO_2 concentrations in different ways - i.e. relationship of CO_2 assimilation and transpiration loss, phenological differences etc.). A clear attribution is not possible but we improved the sentence to avoid confusing the readers.

P. 8 line 35: ". . . between PFTs that might otherwise be lost . . ."
Corrected

P. 10, line 19: "The surface runoff simulated here was found to be consistent with expected patterns" – This is not really true, since no validation of runoff has been made.
We agree with the reviewers. Due to a lack of available data for a thorough large-scale comparison we only provide a descriptive evaluation. However, as we plan a future use of FPC and runoff in a coupled-model we still wanted to include runoff results here for completeness. We changed the sentence to: "The surface run-off simulated here was found to be consistent with the general expected patterns, although a thorough analysis was not possible and will be included in future work." In general we added caution notes in the manuscript whenever we discuss runoff results to make the reader aware of this.

P. 12, line 38: Hickler et al., 2015; Zhu et al., 2016 – Please refer to earlier literature, this has been discussed much earlier by many authors.
We added the classic article by Farquhar et al., 1980 and removed Zhu et al. 2016 to keep the reference list reasonable (we kept Hickler et al., 2015 as it provides a nice summary of the current view on the effect of atmospheric CO_2 levels on plant physiology based on observations and the implementation in vegetation models).

Table 1. Please provide, as far as possible, example species for all PFTs
We extended the table with a comprehensive list of example species.

Table 2. It might be interesting to also list in this table the PD biome areal extent from the observed map of figure 1, in order to compare them with the model
While we agree with the reviewer that this addition would be interesting, we did not include this column as not all biome classes of the model setup align with the simplified biome classification presented in Fig. 1. The biome map of Fig.1 was generated by aggregating floristic classification units of Luebert and Pliscoff, 2017. However, the scheme does not allow to easily discriminate for instance between Matorral and Sclerophyllous Woodland. Furthermore, these classifications often list complex topographic units that do not easily translate into the biome system used.

Table 3. According to the legend, the table lists PD absolute values, but relative changes (in % ??) with respect to PD for the LGM and MH. However, the title at the top of the table, runs over three columns, which is misleading, because it suggests that all values are % cover or mm yr-1, i.e., absolute changes. Please revise.
We agree with the reviewer that this was potentially confusing for the readers and added the suggested improvement (column-specific units) to the table.

Figure 1. It might be useful to provide a map of elevation next to the vegetation map

We changed Figure 1 as requested and added the elevation map as a second plot panel (now Fig 1. a, b).

Figure 8. Legend, line 11: “dark grey” instead of “darkgrey”
Corrected

Appendix A. P. 32, line 8: “... completely different shapes...” instead of “...completely shapes.. .”
Corrected