General response to both referees:

We thank the referees for their helpful and valuable comments, which improved our manuscript significantly. In the following we highlight the major changes and re-interpretations we have made, we added point-by-point answers to all referees' comments and the revised Tables and Figures.

1) Both referees indicated problems with the use of citations in some sections of the manuscript, which we tried to correct.
2) Both referees indicated insufficient justification of our initially stated hypotheses; hence we tried to add a better framing of our study. Additionally, we tried to improve our hypotheses adapted the discussion section accordingly.
3) As suggested by referee 2, we are now able to demonstrate that increasing NH4+ concentrations in the soil water solution could have induced a decrease of amoA genes at the more intensively managed meadow. At the abandoned site, in turn, the higher amount of soil organic matter appeared to buffer decreases of soil water during drought periods, thus ammonia concentrations rose significantly, however not to the same extent as at the meadow. Therefore archaeal amoA gene abundances might not have been as affected by drought as at the managed site. To show this we added a further figure (new Figure 3) and the according statistical analyses in Table 1 and 3, as well as a paragraph in the result and discussion section

Response to Anonymous Referee #1

Referee: In the submitted manuscript, the authors presented results from an in situ experiment of simulated drought in two grasslands experiencing different intensity of land-use and its impact on N cycle key processes and AO. The information presented here is of general good interest and is valuable information. The experimental set-up looks 'clean' and of good quality. Results are generally presented in a pretty clear and interesting way.
However, as I develop below, I have two major concerns that would prevent its publication as it is:
1) the too frequent random and incorrect use of references for justifying (sometimes wrong) assumptions and

Response: We thoroughly revised the references and hope that their use is now correct.

2) the lack of strength and clarity of the work hypotheses, which to some extent appear as if they had been formulated after the experiment was conducted. Also, I think that the 'recovery' (or not) of the tested parameters after the drought ended should be more discussed.

Response: We revised and rewrote the hypothesis section, also in accordance with Referee #2, to add more clarity and emphasis on our research question in the introduction, as well as in the discussion section (see comments below). Moreover, we discussed the dynamics in the recovery phase after drought in more detail.

Major points:

Referee:

1) Here are some examples of 'mis-citations':
Page 9186, Line 17-22 'archaeal AMO has a higher affinity for ammonia and they seem to have a clear advantage in environments with low ammonia concentrations (Gubry-Rangin et al., 2011; Höfferle et al.,2010; Offre et al., 2009; Tourna et al., 2008, 2011)'. Gubry-Rangin et al. 2011 study the niche specialization of AOA lineages, while Offre et al. study the inhibition of AOA growth by

acetylene. None of these references mention AOB or compare AOA and AOB in terms of their affinity for ammonia or 'advantage in environments with low ammonia concentration.
Page 9187, Line 22-26 'Moreover, AOB and AOA seem to differ in their sensitivities to changes in soil water availability (Gleeson et al., 2010; Stres et al., 2008; Szukics et al., 2012), with growth of AOB, but not of AOA, being favoured at higher levels of soil water content (Bates et al., 2010; Szukics et al., 2012).' Page 9197, L10-12: 'Abundances of bacterial and archaeal ammonia-oxidizers have been shown to strongly differ in soil NH4+ concentration optima (Gubry-Rangin et al., 2010; Offre et al.,2009; Schauss et al., 2009)'
Gleeson et al. show a slightly greater impact of WFPS variation on AOA community structure than AOB community structure and NO difference in terms of abundance.
Stres et al. study total archaeal and bacterial communities targeting 16S.
Response: We carefully revised this paragraph in the introduction section to erase any incorrect use of citations. Moreover we rewrote the paragraph, to better justify our hypotheses and hopefully meet both Referee 1 and 2's points of criticisms.
The newly phrased paragraph:
"In many soils archaeal *amoA* genes are more abundant than their bacterial counterparts (Alves et al., 2013; Leininger et al., 2006; Prosser and Nicol, 2008), but AOB seem to outcompete AOA and dominate nitrification in agricultural soils (Jia and Conrad, 2009), N-rich grasslands (Di et al., 2009), and at high levels of NH_4^+ (Brankatschk et al., 2010; Di et al., 2010; Verhamme et al., 2011). Archaeal AMO, in turn, appears to have a higher affinity, and lower inhibition constant for ammonia (Martens-Habbena et al., 2009; Prosser and Nicol, 2012), which could be advantageous at low ammonia concentrations."

Referee: Szukics et al. did not see significant differences in AOA and AOB numbers (even though the text in this article mention it for one of the soil they tested, standard errors presented in their figure 1 clearly show no difference). If they actually do, it is rather the absence of impact of higher moisture content on AOB abundance while it may have decreased AOA abundance, which is different from what has been stated. Bates et al. does not even study water content or AO!!
Response: We re-structured the paragraph and focused on results of studies on drying and rewetting dynamics, in order not to over-interpret results from water addition studies.
The newly phrased paragraph:
"Whereas dynamics of AOB and AOA to ammonia are well studied, much less is known about responses of ammonia-oxidizers, specifically of AOA, to drought and rewetting under *in situ* conditions. Generally, AOB and AOA feature different physiological prerequisites (Schauss et al., 2009), presumably also leading to different responses of AOB and AOA to soil drying and rewetting. …"

Referee: The correct and appropriate use of references should imperatively be thoroughly revised throughout the article to give them back their original meaning and relevance.
Response: We carefully revised the manuscript and accordingly changed parts of the introduction to make it clearer. Moreover, we replaced or added some references (see earlier comment), also in response to Referee #2, to add more power to our hypotheses.

Referee:
2) The first hypothesis is valid, although this should be reminded here that this is because mineralisation is a 'broad' ability while nitrification is a 'narrow' one, if I understood properly the

justification of the hypothesis. On the other hand, there is no justification for the second hypothesis in the Introduction section, just the above-mentioned list of citations, which were misused and/or over-interpreted. The rationale behind the 3rd hypothesis is not clearer or stronger: land-use may have an impact on the parameter you state, but what would be the link with resistance to drought, what would be the underpinning mechanism? And you don't precise what impact, just 'stronger impact'. On what?

Response: We agree that the second and third hypothesis were not very clear and, also in accordance with referee #2, changed them. We re-phrased parts of the introduction to emphasize the rationale behind our new hypothesis.

The newly phrased paragraph:

"…We hypothesized that the phylogenetically 'broad' process of N mineralization is less affected by drought than the more 'narrow' process of nitrification. We expected AOB and AOA to respond differentially to drought; more specifically, that archaeal *amoA* abundances will decrease as a consequence of rising ammonium concentrations in the soil solution. Additionally, we hypothesized that the impact of drought on N-turnover and ammonia-oxidizer abundances will be stronger on the managed meadow than on the abandoned site, as the higher soil organic matter content at the abandoned site could act as buffer against soil drying (Brady and Weil, 2002; Franzluebbers et al., 2002)."

Minor points:

Referee: Page 9186, L17-20: convoluted wording. Please rephrase.

Response: The paragraph was rephrased, as requested.

The newly phased paragraph:

"These individual steps in soil N cycling occur on different temporal and spatial scales (Schimel and Bennet, 2004) and have been shown to strongly differ in their response, both direction and magnitude, to drought (e.g., Auyeung et al., 2012; Chen et al., 2011; Emmett et al., 2004; Gleeson et al., 2010; Stark and Firestone, 1996)."

Referee: Page 9188,L18: Please state why this missing information would be important to know.

Response: A clarifying statement was added:

"However, well water-supplied ecosystems, such as many grasslands in mountainous areas, will be experiencing a higher frequency of drought and heavy rainfall events (Gobiet et al., 2013; IPCC 2012). These grasslands, which are often in transitions from land-management to abandonment, play a pivotal role in nutrient retention and erosion protection with repercussions on densely populated watersheds downstream. Thus, there is an urgent need study possible effects of such climate extremes on soil N cycling in situ."

Referee: Page9189, L21: why 'respectively'?

Response: at both sites (the meadow and the abandoned site) we set up four drought and four control plots; for more clarity we removed "respectively".

Referee: Page 9190, L10: why pooling 2 subsamples? Pooling imply that you're aware that spatial heterogeneity might be important. If so, why just 2 subsamples?

Response: In field studies a certain degree of spatial heterogeneity cannot be avoided. However, because of the spatial limitation of the rain-excluded area we have chosen to pool only two subsamples per replicate.

Referee: Page 9191, L14-15: In any case, this will always give you the potential rates only and not the actual processes, whatever the volume of water you add! It is still valuable information per se.
Response: We agree.

Referee: Page 9192, L14: Strange dilution, but why not! In any case, you should rather state the quantity of DNA that was added per reaction.
Response: The amount of DNA ranged between 2 and 5 ng µl-1. The information was added in the method section**.**

Referee: Page 9192, L15: Were the plasmid used as standards linearized? Close plasmids may have a strong influence on qPCR efficiency and reliability.
Response: Since bacteria and archaea commonly have circular genomes, we decided to use the PCR ™ 2.1 Vector of lifetechnologies, which is a very common circular plasmid for DNA standards in quantitative PCR studies. Of course the topic of using linear or circular plasmid DNA standards is controversial discussed. However, recent studies have shown, that the over-estimation of circular plasmid DNA standards is more pronounced in eukaryotic than in archaeal or bacterial systems. For example, Oldham and Duncan (2012) showed that the ratio of estimated to predicted 16S rRNA gene copies ranged from 0.5 to 2.2-fold in bacterial systems and 0.5 to 1.0-fold in archaeal systems, while the circular plasmid DNA standards grossly overestimated the numbers of a target gene by as much as 8-fold in an eukaryotic system.

Referee: Page 9192, L18-19: There is no need for efficiency formula, but r-squared information would be useful.
Response: for AOA and AOB r-square of qPCR efficiency were 0.997 and 0.999, respectively. The information was added in the method section**.**

Referee: Figure 1: I think that the way you indicate your stat results is a bit complicated and that the figure lacks of visibility. Maybe colours would help?
Response: Figure 1 differs from the others figures as we explicitly wanted to show the temporal variability over the course of the rain-exclusion experiment. We tried to add colours, but this did not increase the readability of the figure and we believe that it is clear enough to understand.

Referee: Figure 2: Would it be possible to indicate the stat results on these barplots?
Response: We added summaries of results of the two-way ANOVA in the barplot-graphs (Fig 2-5). However, paired t-tests were conducted to evaluate differences between control and drought plots on different samplings, but results were only added when differences (after the very restrictive Bonferroni correction) were significant at a $p<0.05$ level.

Referee: Paragraph 3.2 of the Results section and Discussion section: please keep the use of 'potential' when talking about the process measurements that you've performed.
Response: This was changed accordingly in the results and the discussion section.

Page 9196, L6-7: your hypothesis was not 'distinct response' but 'stronger impact' in the meadow, which is not obvious here since potential nitrate immobilisation and potential nitrification seem to be more affected in the non-managed soil.

Response: This sentence was removed, as we re-wrote the hypothesis (see earlier comment).

Referee: Page 9196, L8-12: Okay, it sounds to make sense. But are you suggesting that the difference between the two sites could be due to mowing? It would be risky since you have not tested this (there was no 'mowed abandoned site' or 'unmowed meadow': : :
Response: In accordance with Referee #2 we removed speculations about responses to mowing from our manuscript.

Referee: Page 9196, L14: Why 'however'? What is the link between the two sentences?
Response: We re-structured the first paragraph of the discussion for more clarity.

Referee: Page 9197, L3: Yes for Zhang et al., not for Gleeson! In any case, do you have any hypothesis about why AOB or AOA may be more responsive to drought stress? Maybe related to the concentration of ammonium in the soil solution with various moisture?
Response: We agree and added the suggestion of the referee**.** In our newly phrased hypothesis we now state that during drought, ammonium concentrations in the soil solution increase, which could induce different responses of AOB and AOA. To support this, we added additional information about the calculated concentrations of EON, NH_4^+ and NO_3^- in the soil solution in the results (Table 1 and 3, new Figure 3), as well in the discussion.

Referee: Page 9197, L5-7: again here, why Bates??? Also, Szukics may observe higher AOA abundance at their lower water content, in only one of their soil, but it was actually normal moisture for this soil as they state in their M&M (it's WFPS and not WHC!).
Response: We removed the two citations.

Response to Anonymous Referee #2
Referee: The manuscript titled "Effects of drought on nitrogen turnover and abundances of ammonia-oxidizers in mountain grassland" presents an interesting manipulation whereby two alpine grasslands under different management strategies are subjected to drought conditions in situ in order to observe effects on key N cycling processes related to the generation and consumption of reactive N. Overall I found the approach to be innovative, the methods to be well executed, and the results to be of interest to the readership of Biogeosciences if couched in the proper context. However, similar to reviewer 1, I found some fundamental flaws in the justification for their hypotheses. The incorrect use of references aside, the hypotheses are completely lacking in substance and provide no mechanistic framework to build upon the ideas presented in the introduction. Moreover there seems to be a bit of contradiction in the formulation of the authors' arguments.

Referee: Form previous work, we understand that the kinetics of AOA activity indicates a higher affinity for ammonia as demonstrated by a half saturation constant in the nM range, and the authors even acknowledge that AOA should have a clear advantage in environments with low ammonia concentrations. Therefore, under drought conditions, we would expect ammonia concentrations in pore water to increase over time. Yet, the authors predict that drought would have 'stronger effects on bacterial than archaeal amoA gene copy numbers'. If the authors had alluded to the hypothetical mixotrophic nature of AOA (sensu Jia and Conrad, 2009, or Sims et al. 2012), then perhaps this hypothesis would make sense; AOA weathers drought more effectively than AOB by employing

alternative metabolic pathways. Granted I don't believe this to be the case, but perhaps the idea has a better foundation than the one provided in this manuscript. The authors' would do well to revise their hypotheses for clarity, so the reader better understands their reasoning for why drought affected AOA communities should outperform AOB, or what exactly from a land management perspective is meant by 'stronger impact'. As an example of my meaning, I draw the authors' attention to a recent publication on the same subject matter (doi: 10.1111/1574-6941.12395). In their presenting their hypotheses, Thion and Prosser (2014) provided a clear rational as to why they believe a particular group within the AO community would perform better under drought conditions. I think the authors should give additional consideration to Thion and Prosser (2014), because like their own study, this research was conducted on non-adapted AO communities, meaning that microbes in these soils rarely experience drought. Some of the literature cited in the introduction as a 'case in point' stems from research in Mediterranean climates, e.g. California grasslands, where microorganisms are well-adapted to seasonal drought.
Response: We are thankful for the helpful comment and considered the suggested paper by Thion and Prosser (2014) in revising our hypothesis on effects of drought on bacterial and archaeal amoA gene abundances. We rewrote the respective parts of the introduction and also revised our second hypotheses accordingly. We also agree that soil microbial communities could respond differently to drought, based on the level of pre-adaption or drought history.

Referee: I think the approach of using rainout shelters rather manipulating soil moisture in mesocosms was a good choice; however, it exposed the desired treatment effect to number of modulating factors likely related in part to differences among vegetation communities (primarily grassland vs. grassland populated with ericaceous shrubs, and a legume), e.g. litter production, root density, etc. This may have robbed the study of some of its statistical inference, and perhaps a power analysis a priori would have helped in improving the design. I find it interesting that in choosing this design, the authors focused solely on microbial dynamics and gave no consideration to plant contributions (e.g., N uptake preference, timing of maximal root growth, etc.) to N cycling dynamics in their discussion, despite that plants contribute substantially to below ground processes. Presumably, the plant component, unlike the prokaryotic component, straddles the treatment effect (rainout shelter). Perhaps some of the spikiness in pool dynamics among control plots could be related to differences in plant phenology.
Response: We agree that experiments under field conditions do not have the power to elucidate the mechanisms behind effects of drought and subsequent rewetting on N transformation processes. However, this study rather intended to understand the effect of drought at the ecosystem level, necessitating an integrative approach.

Minor points:
Referee: How does measuring abundance of AOB vs. AOA really get at the functionality of AO in response to drought, particularly from a climate change perspective, since these data may provide a framework for process models? Several studies indicate that, at least in certain environments, population abundance alone is a poor predictor of relevance to ammonia oxidation.
Response: We are aware that gene abundances are only one out of many measures to address responses to changing environmental conditions, thus we tried not to over-interpret our results and addressed this topic in the introduction, as well as in the discussion. Moreover, in our study we combined gene abundances with measurements of microbial gross N-processing rates, which could be included in process models.

Referee: Ctot and Ntot were determined on an EA coupled to an IRMS. Why then not report isotopic values for these elements in Table 1? Some of the readership might draw inference from these values.
Response: We added information on the isotopic composition in Table 1.

Referee: How was the efficiency for the 15N microdiffusion determined, and why is it not reported?
Response: The 15N recovery efficiency of the method was reported to be reproducibly high (Lachouani et al. 2010, Sorensen and Jensen, 1991) but it was not specifically determined in this experiment.

Referee: Regarding linearization of plasmid DNA used to standardize qPCR, I agree with reviewer 1, but here are two publications presenting mixed results to help you make your own determination.
1. Oldham AL, Duncan KE (2012) Similar Gene Estimates from Circular and Linear Standards in Quantitative PCR Analyses Using the Prokaryotic 16S rRNA Gene as a Model. PLoS ONE 7(12): e51931.doi:10.1371/journal.pone.0051931
2. Hou Y, Zhang H, Miranda L, Lin S (2010) Serious Overestimation in Quantitative PCR by Circular (Supercoiled) Plasmid Standard: Microalgal pcna as the Model Gene. PLoS ONE 5(3): e9545. doi:10.1371/journal.pone.0009545
Response: We acknowledge differences between plasmids (see earlier comment to Referee #1).

Referee: Throughout, either/or and neither/nor.
Response: this was corrected accordingly throughout the manuscript

Referee: Throughout, please note archaeal, not archael
Response: this was corrected accordingly throughout the manuscript

Referee: In the Discussion, comments concerning mowing effects are overly speculative and should be removed.
Response: also in accordance with Referee #1 we removed speculative comments about mowing effects.

Referee: 'in accordance with' not 'in accordance to'
Response: this was corrected accordingly

Referee: 'respond more sensitively' not 'more sensitive'
Response: this was corrected accordingly

Referee: P9197 L3-9 Clunky and redundant, please consider revising.
Response: Because of the new structure in parts of the discussion section this sentence was removed.

Referee: Soil acidity (_5.5) may play a role in overall nitrification potential, particularly with regard to competitiveness of AOA to AOB. Why was this never really discussed?

Response: We agree that the soil pH may play a role in the overall nitrification potential. The soil pH was however not differing between sites. Thus we expected that pH might not be a determinant for different responses of AOA and AOB to drought.

Referee: Pg 9197 L17-21 Please clarify, AOA and AOB were extracted to determine nitrification potential? Are the authors referring to abundance as a potential? If so please refer to the first comment in this section.
Response: No, we did not refer to gene copy abundance as a nitrification potential, (see answer to the first comment). The sentence was rephrased for more clarity.
"It might also hint to functional and structural differences of AOA (Alves et al., 2013), as well as of AOB populations between the studied sites (Gleeson et al., 2010)."

References:
Lachouani P, Frank AH and Wanek W: A suite of sensitive chemical methods to determine the d15N of ammonium, nitrate and total dissolved N in soil extracts, Rapid Communications in Mass Spectronomy, 24, 3615-3623, doi:10.1002/rcm.4798, 2010.

Oldham AL, and Duncan KE: Similar Gene Estimates from Circular and Linear Standards in Quantitative PCR Analyses Using the Prokaryotic 16S rRNA Gene as a Model. PLoS ONE 7(12): e51931.doi:10.1371/journal.pone.0051931, 2012

Soerensen P and Jensen ES: Sequential diffusion of ammonium and nitrate from soil extracts to a polytetrafluoroethylene trap for 15N determination. Anal.Chim.Acta 252, 201-203, 1991

Thion C and Prosser JI: Differential response of non-adapted ammonia oxidising archaea and bacteria to drying rewetting stress, FEMS Microbiology Ecology, doi:10.1111/1574-6941.12395, 2014