Reviewer 2

Comment 1

There were 18 plots, and there were multiple samplings per plot, and at multiple times. Obviously, samples taken within the same plot are not independent of each other - they are closer to each other than they are to samples in different treatments, and are likely to have similar values. The appropriate thing to do would be to use their mean values per plot.

 \rightarrow The tests were run again using mean values per plot.

No major changes occurred in the outcome of the tests and no changes had to be made in the interpretation of the results.

Changes included:

Material and methods

Line 208-209:

Differences in ecosystem N stocks, biomass and C stocks were tested with **two-way ANOVA's (instead of generalized linear models)** in the R software (R Development Core Team, 2012), with seabird colony (yes – no) and tephra depth (deep-shallow) as fixed variables.

Line 211-212:

Deleted: Plots and sampling year (2012 and 2013) were used as random variables. Random factors were excluded whenever insignificant.

Results

Line 243-245:

As there was no significant difference in N accumulation rate between the shallow and deep tephra sand inside the seagull colony, the estimated N accumulation rates of 36 and 58 kg ha⁻¹ y⁻¹ on shallow and deep tephra sand respectively, was changed into an estimate of 36 - 58 kg ha⁻¹ y⁻¹ inside the colony.

Line 257:

As there was no effect of tephra depth on total C stocks inside the seabird colony, the sentence that **deep tephra sand stored more total C than shallow tephra sand** was **deleted.**

Discussion

Line 335:

As there was no significant difference in N accumulation rate between the shallow and deep tephra sand inside the seagull colony, the input rate of 47 kg ha⁻¹ y⁻¹ averaged over both substrate types was changed into an input rate of on average 47 kg ha⁻¹ y⁻¹.

Using mean values per plot had no impact on the discussion or conclusions of the article.

Comment 2

I would ask them for an explicit model, complete with degrees of freedom associated with the F-statistics (numerator and denominator).

 \rightarrow The model that we used was a two-way ANOVA. We tested the effect of 'colony' (fixed variable) and 'tephra depth' (fixed variable) on the dependent variables ((1) N stocks, (2)

Authors comments to the reviewers questions – Iteration: Major revision

Biomass stocks, (3) C stocks, (4) N accumulation rate, (5) SON per soil depth and (6) root N per soil depth).

First, the interaction between 'colony' and 'tephra depth' was tested. When there was no interaction, the test was split up in two one-way ANOVA's to test the separate effects of 'colony' and 'tephra depth'. In case of interaction, four one-way ANOVA's were conducted, one for each combination of treatments.

The complete list of degrees of freedom (numerator and denominator) and F-statistics can be found in Table 1 and 2.

Comment 3

I would also not have thought as "year" as a random variable; it would make more sense to explicitly test whether year has an impact, since that would certainly change the interpretation of results.

→ Year was not included in the tests, as we measured different plots for each treatment in the different years (instead of taking replicate measurements in the same plots). This experimental design was chosen in order to obtain sufficient replicates for each treatment during the very restricted time period that research is allowed on the island. We are convinced that this combination of years does not pose a major problem, because stocks vary only little from year to year. Of course, this also implies that it is not possible to distinguish the year effect from the plot effect.