We would like to thank both Reviewers for their careful and constructive reviews of our paper. In response to their suggestions, we did the following main changes to the manuscript:

1) we added a figure showing the monitoring network and the surface topography of the nine analysed glaciers (Figure 2)
2) we modified the last figure (Figure 8), adding the cumulated residuals, as suggested by the Reviewer 2
3) we modified the structure as suggested by the Reviewer 1, adding a Section with the available datasets, separated from the Methods
4) we added error estimates in mass balance measurements (Section 3.1)
5) change points in time series have been identified using specific statistical tests
6) we added a discussion on the spatial representativeness of analysed glaciers
7) we quantified the relationship between seasonal and annual NAO and the seasonal and annual components of the mass balance, comparing the results from the Italian glaciers with those from other nations in the European Alps. A table was added reporting the statistics, and a discussion was added in section 4.2
8) the references suggested by the Reviewers were added to the reference list
9) we implemented the suggested changes and answered to the specific comments made by the reviewers, as detailed in the following of this document. The comments are reproduced in italic while the authors responses are reported right below. Line and page numbers are referred to the discussion paper.

Reviewer 1

R1 – 1: This paper presents a comprehensive analysis of the mass balance data acquired on all Italian glaciers since the beginning of the measurements. The authors compile data covering one to five decades that mostly have a seasonal resolution and interpret changes in mass balance in connection to climatic forcing (temperature and precipitation). Data from long-term monitoring programmes are often difficult to be published and performing innovative science based on them is challenging. The authors do a good job in exploiting the full data basis for Italian glacier mass balance to come up with sound conclusions and reasonable recommendations for the future. Nevertheless, I somewhat miss the "breaking news" in this study. Many of the findings are not actually new and the authors can often just confirm what has been said earlier. Of course, this is a problem which is not trivial to be addressed provided the availabledata that are discussed. However, by slightly shifting the focus to the more exiting and novel results I have the impression that this paper could considerably gain in quality and scope. This would mean shortening the article in certain places and providing more detailed analysis at others.

In our opinion, the value of this paper lies in the joined presentation and analysis of the mass balance data from the Italian glaciers, which had not yet been carried out. We tried improving the paper as suggested, focusing on the parts recommended by the two reviewers.

R1 – 2: I see several additional important and interesting aspects that could at least be partly addressed by the authors: Data quality: An integrated interpretation of a data set as the one presented in this paper crucially relies on the quality of the data. At present, however, only very little is said about this point and no specific analyses are performed (uncertainty of the data based on general literature values). As much as I know, no homogenization (e.g. by comparing geodetic and glaciological mass balances) has been performed *for most of the Italian glaciers. To ultimately increase the value of a data set, assessing the quality and performing a homogenization would be of utmost importance and should actually be performed before interpreting the data in a climatological context. I fully understand that this is not feasible in the frame of this study but the authors might still consider to dedicate some discussion or even preliminary analyses (e.g. based on published results) regarding this aspect*.

Published accuracy estimates and available comparisons between the geodetic and direct methods have been added in section '3.1 Mass balance measurements and calculations'.

R1 – 3: Information on investigated glaciers: Related to my point above, I also missed more detailed information on the sites. There are only tiny maps showing the individual glaciers that do neither contain contour lines nor the distribution of the measurement network. For such an overview study, it would be important to present the monitoring strategy so that the reader can judge the context of the data.

A new figure was added, reporting the measurement network and the contour lines for each of the nine analysed glaciers (Figure 2).

R1 – 4: Representativeness of glaciers: In my opinion, a central question of such an integrated analysis would be the representativeness of the individual series for a larger region (e.g. the Italian Alps). This would be very valuable for other scientists looking for time series that well describe the spatio-temporal mass balance variability and are not biased by local effects (such as Careseer, see Carturan et al., 2013). Unfortunately, this point is only marginally addressed in the paper. Given the approaches presented here, this aspect would be possible however to be investigated in more detail. This might allow some useful recommendations for the glaciological community.

Considerations on the spatial representativeness have been added in Section 4.1, based on available geodetic mass balance calculations.

R1 – 5: Structure: The description of the data and the study sites is placed in the Methods section, which is not ideal. These chapters definitely need to be separated, especially for a paper that principally relies on a strong data basis.

Ok, modified as suggested.

R1 – 6: Detailed comments: Page 5853, line 20: Without knowing about the distribution of the measurements (figure) the point density is difficult to be interpreted

Please, see the reply to comment R1 – 3

R1 – 7: Page 5855, line 10: Identical methods used for temperature and precipitation extrapolation? Partly unclear.

Modified to improve clarity

R1 – 8: Page 5855, line 29: Several studies have already indicated that at least some of the nine WGMS reference glaciers are not (not anymore) actually representative for the European Alps. This might influence comparisons as the one presented here.

A sentence was added in Section 4.1 to highlight this aspect

R1 – 9: Page 5861, line 20: When discussing potential NAO effects it would be important, in my opinion, to be more quantitative: The present statements are fully qualitative. If really entering this topic the authors should be more specific and provide statistical numbers and evaluations in the context of their data.

We analysed the correlation between the NAO and the seasonal/annual components of the mass balance. Please, see the reply to the comment R2 – 25.

Reviewer 2

R2 – 1: General comment. L. Carturan and co-authors provide analysis of glacier-wide mass balance times-series for 9 glaciers in the Italian Alps covering a period of record from 10 to 47 years. The link of the mass balance variations to temperature and precipitation chronicles is then investigated. Annual balances are reasonably reported to be highly connected to winter (October-to-May) precipitations and summer (June-to-September) temperatures. For the longest available records, the increase in mass loss observed since the 1980's is related to the warming and lengthening of the ablation season. Changes in correlations between mass balances and summer temperatures/winter precipitations is interpreted as possible milder accumulation seasons in the last years of the time-series, but the changes in glacier topography in response to climate cannot be ruled out.
The main value of this paper is to gather the longest mass balance records available in the Italian Alps and to propose a joint analysis. Moreover, I find the examination of the changes in correlation interesting as climate cannot be assumed in steady state in the time series analysis. Nevertheless, the authors hardly achieve in drawing new and firm conclusions on how the regional and larger-scale climate drives glacier mass changes. Despite the reported material is instructive and of quality, the paper is not enough innovative, and below and in my substantive comments I have suggested some way to strengthen it. I have three key suggestions below.
Despite mountain glaciers being recognised as excellent indicators of climate change, this paper demonstrates once again that inferring the climatic signal from a glacier mass balance series is not an easy task:
1°) A first reason is the intrinsic limit of glacier-wide balance series which includes both the climate signal and the effect of the changing glacier topography in response to climate. A time signal free from geometry changes and site effects should be first extracted from the series before investigating the potential link with the climate.
2°) Secondly, the time-structure of the extracted temporal signal should be analysed quantitatively to detect trends, change point in mean and variance with appropriate empirical statistical tests, or modern and advanced treatment like Bayesian inference.
3°) Third, the authors should take advantage of the stratigraphic-floating date system to analyse how the ablation duration may have changed from year–to–year; and subsequently derive the rates in ablation which is a way to quantify surface energy fluxes responsible for melt.
Different issues can be addressed to extract the climate signal:
- The issue raised by Elsberg and others (2001) of computing a mass balance referred to a constant topography. This requires the time change of the area and elevation of the glacier surface.
- Use point balances as adopted by Huss and Bauder (2009), or Vincent et al., (2004) which removes the effect of the surface change but still requires accounting for the lowering of the surface elevation of the glacier.

- Adopt a variance decomposition to separate the overall annual effect from the spatial variability at the glacier surface (Lliboutry, 1974; Rasmussen 2004; Eckert et al., 2011); this requires also accounting for elevation changes.

We agree with the Reviewer that part of the glacier response in terms of annual mass balance depends on the geometric adjustments. The methods for extracting the climatic signal, however, require information on the area and elevation changes, and/or the availability of point mass balances, in the observation period, which unfortunately are not available for most of the glaciers analysed in this study. The change points in the time series of mass balance have been extracted using statistical test, as suggested (see the reply to comment R2 – 15 for details). The dates of the surveys in the floating-date system do not provide useful information on ablation duration, because they have been selected based on weather conditions, accessibility of the glaciers, availability of personnel, and they did not necessarily correspond to mass balance minima.

R2 – 2: Points/limitations of minor extent are: Meteorological series are reconstructed locally from a linear composition of climatologies (constant spatial fields) and anomalies (uniform temporal deviations). Implicit is that spatial fields are stable in time and that the climate is in steady state which is a strong hypothesis.

Spatial fields in the reconstructed dataset are not stable in time. In this case only single points referred to the monitoring sites were reconstructed, but if a high resolution grid were reconstructed for e.g. the last 60 years, the result would be that spatial gradients are not constant through time because also temperature anomalies (year-to-year variability) and the long-term tendencies (trends) are characterized by a spatial gradient captured by the interpolation of the temporal component. This make the spatial gradients in climate normals not constant in time. As an example let us suppose that the vertical gradient in temperature climatology is of -0.65°C/km in 1960s and that from 1960 to 2010 there is a temperature trend which is strongest the highest is the elevation. In this case, if we extract the climate normal for the decade 2010s we will observe a vertical gradient which will be reduced with respect to the -0.65 observed in 1960s. The same is true in the other x and y components.

R2 – 3: Time-series in glacier-wide balance should be controlled and if necessary homogenized by the geodetic method (photogrammetry, laser scan or lidar altimetry) to discard any bias/trend that may corrupt the time signal and its link to climate.

The available comparisons between the geodetic and the direct methods have been added in Section 3.1. For the analysed glaciers, these assessments reveal that no calibration of direct mass balance results is required.

R2 – 4: Quantifying the link with NAO for the selected glaciers will be innovative.

We have analysed the correlation of the seasonal and annual mass balance components from the Italian glaciers and the NAO, describing the results in Section 4.2 and reporting them in Table 5 (see the reply to comment R2 – 25).

R2 – 5: There follow several detailed questions, comments, suggestions, and indications of minor flaws in the paper. Substantive comments. P5850-L22. This is true for glaciers at the melting point. For cold and polythermal glaciers, the atmospheric conditions (atm. temperature and precipitations) control the thermal regime (Hoelzle et al. 2011; Gilbert et al., 2012), and the mass balances of those glaciers are not directly connected to climate (see e.g. Vincent et al. 2007).

We agree that atmospheric conditions control the thermal regime of cold and polythermal glaciers, and that these glaciers (or upper glacier parts in the mentioned examples) respond in a different way to the same climatic forcing, compared to glaciers at the melting point. However, the mass balance of cold and polythermal glaciers is actually connected to climate (although in a different manner), and we prefer keeping the sentence as it is.

R2 – 6: P5851-L5. Use here and elsewhere in the paper the terminology "glacier-wide balance" as proposed by Cogley et al. (2011) to whom you refer appropriately.

Ok, corrected.

R2 – 7: P5851-L17. I would propose to add at the end of the sentence "reviewed and analysed jointly."

Ok, added.

R2 – 8: P5851-L22. It is not clear to me what non-linearity means here. Should the link between climate and glacier mass changes be a linear relationship? Is it rather meant that you are in search of a rupture (in mean or variance) in the mass balance time-series?

We meant feedbacks, corrected.

R2 – 9: P5852-L6. Have the mass balance records been controlled by the geodetic method and eventually calibrated as recommended by Zemp et al. (2013)?

This was done for three glaciers and the results indicate that calibration is not required. A sentence was added at the end of Section 3.1

R2 – 10: P5854-L13. In search of the max/min for the winter and annual balances, the time system in which the mass balances are measured rather refers to the stratigraphic system. An important implication of this floating-date system is the opportunity to retrieve summer ablation duration and rates, and analyse how these may have changed from year to year along the record.

Please, see reply to the coment R2 – 1.

R2 – 11: P5854-L16. Regarding error estimations of glacier-wide balance measurements, if not based on a throughout analysis for each mass balance record, I would recommend to use the value of 340 mm w.e. yr-1 provided by Zemp et al. (2013) from a data set of 14 glaciers in Europe.

In the revised manuscript we have reported the range of errors indicated by the investigators in charge of the measurements (end of Section 3.1)

R2 – 12: P5854-L19 -to- P5855-L19. I understand that this method is basically a linear variance decomposition between space and time effects in weather data. This is a strong hypothesis at it assumes a steady state climate, i.e., steady spatial gradients along the analysed period. It would be interesting to validate this assumption at, at least, a site in the vicinity of one of the studied glaciers where a long record is available from a mountain weather station by a comparison with the reconstructed temporal series. Note that further in the paper, only correlations are in search in the time deviation signal. Then if the absolute (local mean) of the reconstructed time series is biased, it won't affect the correlations.

The method is not a linear variance decomposition and, in particular, it does not assume a steady state climate because spatial gradients change along time, in agreement with the observed spatial gradients characterizing both long-term trends and year to year variability. The adopted interpolation methodology is widely used, documented and validated in the scientific literature (see quoted papers in the manuscript).

R2 – 13: P5856-L15. Here the limitation I mentioned in the general comments is critical. Glacier-wide, i.e. surface-integrated (averaged) balance series provided by WGMS are convolutions between climate signal and glacier geometry changes in response to climate (refer to e.g. Elsberg et al., 2001). This results in two main opposite feedback for retreating glaciers that concerns this study: ablation areas of strong negative budget are removed from the average balance, tending to make it less negative. Oppositely, a lowering of the glacier surface tends to force the surface energy balance in relation with altitudinal gradients. How these 2 feedbacks may or not compensate each other depends on each glacier (Huss et al., 2012), but generally contribute to an additional (non climatic) trend in the series. It is not clear how these 2 effects are to be differentiated in the present analysis. I would proposed to use a mass balance referred to a constant topography (Elsberg et al., 2001) or analyse point mass balances (Huss and Bauder, 2009; Vincent et al., 2004) or temporal signals retrieved from them using variance decomposition (Lliboutry, 1974; Rasmussen 2004; Eckert et al., 2011).

Please, see reply to the comment R2 – 1. We are aware of these issues, and discussed them at the end of the Discussion, where we comment the results of the moving correlation analysis between the mass balance series and the climatic variables, reporting the example of the Careser glacier.

R2 – 14: P5856-L18. I would propose to add sub-sections in the 3.1 section to make distinct analyses of Ba, Bw and Bs time series.

In our opinion it is not possible to analyse separately the three components of mass balance and we would prefer to avoid the introduction of additional sub-sections if possible.

R2 – 15: P5856-L23. How do you detect quantitatively the change point in the Ba time series? Do you use a statistical tests such as Mann-Kendall, Mann-Whitney or Pettitt tests?

We detected the change points in the Ba time series using the binary segmentation algorithm (Edwards and Cavalli-Sforza, 1965; Scott and Knott, 1974; Sen and Srivastava, 1975) implemented in the 'Changepoint' R package (Killick and Eckley, 2014). The results obtained with the binary segmentation algorithm were checked using the AMOC (at most one change) method and the segment neighborhood algorithm (Auger and Lawrence, 1989). Sentence modified and reference added in the text.

Auger I.E., Lawrence C.E., 1989. Algorithms for the optimal identification of segment neighborhoods. Bulletin of Mathematical Biology, 51(1), 39-54.
Edwards A.W.F., Cavalli-Sforza L.L., 1965. A method for cluster analysis. Biometrics, 21(2), 362-375.
Killick R., Eckley I.A., 2014. Changepoint: An R package for changepoint analysis. Journal of Statistical Software, 58(3), 1-19. DOI: 10.18637/jss.v058.i03
Scott A.J., Knott M., 1974. A cluster analysis method for grouping means in the analysis of variance. Biometrics, 30(3), 507-512.
Sen A., Srivastava M.S., 1975. On tests for detecting change in mean. The Annals of Statistics, 3(1), 98-108.

R2 – 16: P5857-L13. A winter balance abrupt change point (+23%) has been detected in 1977 in the western Alps (Eckert et al., 2011), and is representative of a regional change in winter precipitations in the French and western Swiss Alps (weather data: Durand et al., 2009ab; glacier winter balances change point 1974-76: Aletschgletscher, Valais-Switzerland, Huss and Bauder, 2009; Sarennes gletscher, France, Thibert et al., 2013;). Do you detect such a change point in your winter mass balance records?

Unfortunately the series of Bw from Careser glacier has some gaps, and therefore it is not possible to calculate change points. However, from the available data (Fig. 2) it can be argued that such an abrut change did not occurred in the area of the Careser Glacier. We have analysed change points in the Oct-May precipitation amounts and we report the results in Section 4.2 (see reply to comment R2 – 23).

R2 – 17: P5857-L15. Even when the AA has disappeared, the winter accumulation still controls the snow-toice transition date and influences the summer balance (due to the albedo feed-back). The way Bs is correlated to Bw at Careser Glacier before/after 1981 may help to quantify this.

We have checked this but the correlation is not statistically significant (r = 0.18 and 0.05 before and after 1981, respectively).

R2 – 18: P5858-L2. "…the decrease in the glacier-average albedo…". Don't you suspect any changes in the debris cover and any albedo change of the ice?

Modified as suggested. The analysed glaciers are mostly debris-free, with the exception of the lower parts of Sforzellina and Ciardoney, and of small marginal parts or medial moraines in the others. However, here the focus is on the decreased albedo from the early disappearance of the snow cover from low-lying and flat glaciers, compared to those which retain accumulation areas in their upper reaches. A short sentence was added concerning te accumulation of dust and debris on the surface.

R2 – 19: P5858-L12. How Bw correlates with October-May precipitations?

The correlations has been analysed and the results are described in Section 4.1.

R2 – 20: P5858-L19. Can you quantify the effect of the lowering of the surface in that Bs trend?

In this period and in Table 2 we compare the mean Ba, Bw, Bs and AAR of the 9 glaciers in the decade from 2004 to 2013. We are not presenting or discussing trends in Bs, and the change in geometry in this period is not available for most analysed glaciers.

R2 – 21: P5860-L3. Are Bw and Bs significantly correlated? I would add a column in Table 4 with the Bw-to-Bs correlation coefficients for each glacier.

For all the analysed glaciers there is no correlation between Bw and Bs. A sentence was added at the end of Section 4.1.

R2 – 22: P5860-L11. As proposed for section 3.1, a divide in 2 sub-sections for temperature and precipitation times series may help to structure and clarify the analysis.

Please, see reply to the comment R2 – 14.

R2 – 23: P5860-L20-22. As proposed for the mass balance time-series, you should base your detection of change point on a statistical test.

We found a change point in 1977, in analogy to the findings of the studies mentioned at comment R2 – 16. A sentence was added in section 4.2, reporting this result.

R2 – 24: P5860-L23. As the measurements are performed in the floating-date system, could winter precipitations be cumulated according to measurement dates at each glacier, and summer temperatures as well?

Please, see reply to the comment R2 – 1.

R2 – 25: P5861-L6-21. The link between mass balance seasonal terms and large scale NAO variable is unfortunately not investigated here but deserves an analysis. How the changes in the different seasonal mass balance terms are controlled by larger-scale synoptic variables remains an open question as contrasted results from bibliography are still hard to interpret. Significant relationships are indeed reported for annual mass balance in Scandinavian glaciers and Svalbard (e.g. Fealy and Sweeney, 2005), with nevertheless substantial differences between northern and southern Scandinavia (Marzeion and Nesje, 2012). For Alpine glaciers, the reported correlation is rather weak for annual balances (Reichert et al., 2001; Six et al., 2001) and decreases from positive to negative as one goes from the eastern to the southwestern Alps (Marzeion and Nesje, 2012). For the French western Alps, Durand et al. (2009b) do not find any correlation between winter/annual NAO anomalies and winter precipitations. Regarding winter balances, Thibert et al. (2013) achieve to the same conclusions at Sarennes glacier (French Alps). Only significant correlations emerge for December to February NAO anomalies, and surprisingly only with the summer balances. It will be interesting to know if this uncommon finding is confirmed or not for the Italian glaciers to see if their mass balance is connected to the NAO through their winter balance or exclusively through the summer balance.

We analysed the correlation between the NAO and the seasonal/annual components of the mass balance, comparing the results from the Italian glaciers with those from glaciers of other countries in the European Alps. The results are reported in a new table (Table 5), and they are discussed in Section 4.2.

R2 – 26: P5862-L1-12. The correlation between October-May precipitations Bw and temperatures should be investigated.

Ok, analysed. The results concerning precipitation have been described in section 4.1 (see reply to the comment R2 – 19). Here we report the only glacier with significant correlation between Bw and October-May temperature (i.e., the Careser Glacier).

R2 – 27: P5863-L3-23. Again here the analysis is limited because of the geometry artefacts that corrupt the glacier-wide Ba series.

We agree and are aware of the limitations related to geometric adjustments. For this reason, we reported considerations about the non-linear response of Ba to atmospheric changes at lines 11-23, showing the emblematic example of the Careser Glacier.

R2 – 28: P5865-L18. The geodetic mass balance should also help to control the glacier-wide Ba series, and construct a constant geometry mass balance record (following Elberg et al., 2001) to be connected to climate drivers.

These considerations have been added at the end of Section 4.3.

R2 – 29: P5866-L1-7. As a principal investigator of mass balance myself, I found that this paper on the longest Italian mass balance series could be an opportunity to associate all persons in charge of the measurements in a common publication.

This work was intended as a first assessment based on published data. We plan to involve the principal investigators in future publications investigating the monitored glaciers in the Italian Alps.

R2 – 30: Figures. Figure 2, 4, 5 and 6 should be enlarged.

We will check appropriate figure sizes when we have the layout.

R2 – 31: Figure 7. The residual sum could also be plotted to highlight deviations along the period.

Ok, the figure has been edited as suggested.

R2 – 32: Stylistic comments. P5850-L12. "precipitations"; "temperatures"

Ok, modified

R2 – 33: P5850-L15. "precipitations"

Ok, modified

R2 – 34: P5850-L16. "temperatures tend to become"

Ok, modified

R2 – 35: P5851-L10. As a first word in the sentence, write "twenty five" in full letters.

Ok, modified

R2 – 36: P5852-L12. "Nine monitored glaciers fulfil these characteristics"

Ok, modified

R2 – 37: P5852-L19. "in the Ortles-Cevedale (Eastern Alps, Fig. 1) since 1967."

Ok, modified

R2 – 38: P5853-L10. "…of repeated readings…"

Ok, modified

R2 – 39: P5853-L24. Spelling of "grand Etrét" has to be homogenized with upper case letter for "Grand" and acute or grave accent for "Etrèt" in text, table and captions.

Ok, corrected

R2 – 40: P5854-L14. "Typical randon errors…"

Ok, modified

R2 – 41: P5854-L15. Again, "glacier-wide"

Ok, modified

R2 – 42: P5856-L19. "The longest available series…"

Ok, corrected