GENERAL COMMENTS

On the whole, I found "Climate suitability estimates offer insight into fundamental revegetation challenges among post-mining rehabilitated landscapes in eastern Australia "a thought-provoking article surrounding how best to target revegetation actions given potential climate constraints. However, I have a number of concerns that I think would require addressing before it is suitable for further publication. In particular, I think the authors need to consider whether NDVI is a good predictor of early establishment which is what they appear to concentrate on. I also think they should consider whether their results are sensitive to different assumptions as outlined in more detail below. Akin to other reviewers, I wonder whether some of the message is obvious but having said that, I think it is good to consider the obvious in a rigorous framework. I believe Audet et al. attempt to do this, and in a broad sense raise some noteworthy points, particularly if mine-site restoration plans don't already consider climate suitability.

On the one hand, Mike Perring's commentary reflects similar concerns to Rev.'s 1 & 2, so (to avoid redundancy, where necessary) we've reiterated responses and amendments to these points below – thereby contributing to a much improved manuscript. On the other hand, he has challenged the use of NDVI as a "predictor of early establishment", which appears to underpin some confusion as to our primary study intentions.

As retorted to Rev.2 (responses #2 & 5), the impetus for creating the rainfall index was to emphasise site sensitivity to climatic factors and to draw attention to further climatic and rainfall metrics (other than simply mean annual rainfall) having potential influence toward land rehabilitation among semi-arid environments. We then determined/compared the relationships of these climatic factors to bioregional vegetation patterns (as estimated by NDVI). This, as opposed to concentrating on NDVI as a predictor of site sensitivity. These points have now been clarified (pg.9, ln.23-27).

Retrospectively, we believe that inclusion of maps for NDVI distribution in eastern Australia (which are important to our synthesis) in the original manuscript could have undermined the primary intentions of the work. And so, we've now relegated and made reference to this figure in the appendix.

SPECIFIC COMMENTS

ABSTRACT I'm not sure that the use of the word 'susceptible' is helpful in this context. It begs the question susceptible to what? I believe it is easier to see it as a suitability index that goes from ideal - moderate - least suitable, with no requirement to use the word susceptible. Later on in the abstract they use the word 'unsuitable' - this communicates the concept perfectly well. Further details on how the suitability index is derived would be good. (I elaborate on this further below).

Refer to response #1 to Rev2. commentary:

Based on these comments and those of the other reviewers, we fully recognise the semantic problems associated with the 'suitability-susceptibility' terms used throughout the original manuscript – this has been outlined in response to Reviewer 1. Consequently, we've carefully amended the terminology by replacing them with 'climate pattern' and 'site sensitivity to climatic factors' while also including precisions to the terminology where required. This has led to necessary changes to the Title, Introduction, Methods, and Discussion sections. Likewise, and where appropriate, we've also amended usage of the terms metrics, estimates and indices throughout as prescribed.

INTRODUCTION Line 13, pg 18547 - not sure why divergence is a result of the land-

form elements. I think this needs elucidating further.
Line 21, ibid - suitability of climate factors. Taking away climate change, the expectation would be that the neighbouring analogues should be suitable references unless other abiotic conditions are radically different) and therefore that the climate would be suitable for these communities. However, this raises the point that what conditions are suitable for establishment may be different to what is suitable for continued maintenance of natural communities (depending on the growth form). This highlights the core of the problem, and I am uncertain how NDVI gets at this core.

Thus, on page 18548, line 5, the authors state: "...analysis seeks to assess various climatic parameters ... that are relevant to rehabilitation development (particularly plant early establishment), and to compare these combined criteria across different geographic locations currently affected by ongoing mining activities". I am unsure how the scale of the NDVI measures relate to the scale of the mine sites, nor how long the mine sites have been undergoing rehabilitation. I think more direct measures of plant early establishment would confirm the suggestions made by the authors - this is mentioned in the discussion but I think needs emphasising more (perhaps in the abstract).

Amended/Statement clarified (pg.2, ln.6-14):

"The (in)ability to achieve an intended rehabilitation outcome is frequently attributed to the radical and potentially irreversible differences between the physicochemical starting point of the post-disturbance environment compared to that of the intended post-rehabilitation outcome (Doley et al., 2012). Adding to these circumstances, an emerging scenario suggests that climatic factors should represent a further overarching challenge toward rehabilitation schemes, particularly in the era of climate change (Harris et al., 2006;Hobbs et al., 2009;Jones et al., 2012). Critical for the success of mined land rehabilitation is the availability of water and hence the climatic characteristic of geographic regions which are defined by a number of weather-bound factors."

Line 25 onwards seems to be a repetition of the abstract.

Amended.

MATERIALS AND METHODS
If BGs policy is for repeatability of work, then I think substantial rewording of the methods is required. I did not always follow the chain of logic and found that some parts obscured rather than illuminated.

This is a critical point (also mentioned by Rev.2 – see response #2 & 3). We've taken care to clarify the description of methods and procedures to facilitate repeatability of the work (pg.9, ln.23-27).

Page 18549, line 13-15 - unclear sentence. line 20 - criteria relevant to the early establishment of native vegetation. Again, does NDVI give a good indication of early establishment? I agree rainfall / soil moisture is clearly important - perhaps some references to papers that demonstrate this in the Australian or worldwide context?

Amended. Refer to response to general commentary.

Line 23 - what is IBRA?
Amended. IBRA has been defined as the Interim Biogeographic Regionalisation for Australia, which is widely used to classify biological regions throughout Australasia.

Line 26 - mean monthly temperature. This is mentioned but then not used in any data

analyses - why?

Indeed, we only focused on rainfall characteristics rather than temperature. We clarified this (which was an error of inattention) in the revised manuscript.

Page 18550, lines 1-5 - state clearly why were the different rainfall metrics chosen (perhaps not just in the Table 2, which is only referred to later in any case). Also justify the different break points for the analysis i.e. >25 and <3mm. Have these been demonstrated as being ecologically and hydrologically important. If so, then just look at that. If not, perhaps look at other values in some kind of sensitivity analysis and see whether this changes the conclusions. Also, why (line 1) are the number of days 'relative'?
Line 7 - "each of the sites' climate parameters was scored qualitatively". How was this done. Were all parameters equally weighted or did e.g. annual rainfall have more weight than days under 3mm? I assume equal weight; would changing weight have any effect on results? Note that some think median rainfall is a better indication of an area's average rainfall than its mean - does this change results?

This is a critical point (also mentioned by Rev.2 – see response #2 & 3). We've taken care to clarify the description of methods and procedures to facilitate repeatability of the work.

Line 7 onwards - this is the area in particular that needs more explanation to be repeatable by others. As well as the points raised immediately above, for % number of days per year with rainfall events above or below a certain threshold - were these calculated out of total number of rainfall event days or total days in a year?

The purpose of the parameters $R_{d,25}$ and $R_{d,3}$ was to represent the occurrence of rainfall events with high and low intensity, respectively. We recognise that metrics are available that reflect on these rainfall events much better. Hence, we've replaced these values with Average Recurrence Intervals (ARI). ARI represents the frequency of selected rainfall events (characterised by rainfall intensity and duration). In the revised manuscript, we selected three extreme rainfall events, all of which deem to be critical for initial ecosystem establishment (pg.5, ln.2-16):

"Site specific values of long-term rainfall parameters and vegetation density are shown in Table 3. From these data, each of the sites' climate parameters was scored qualitatively − i.e., being either less sensitive [ideal], moderately sensitive [adequate], or highly sensitive [potentially problematic] − in relation to a series of rainfall criteria which also included a description of the given parameter's indication of biological significance (Table 2). [...] For example, annual rainfall depth (Rd) and average recurrence interval of prolonged events with low intensity (ARIrehab) were deemed to be indicators of the sites' general level of water availability. Contrary, the average recurrence intervals of short (ARIstorm) and prolonged events with high intensity (ARIcyclo) – representing storm and cyclone events, respectively – indicate problems of erosion or inundation, respectively."

Page 18551, equations. Could justify combination in arithmetic way. Why not multiplicative or a ratio? Does this change results in any way? Page 18552 - line 8 onwards. Could multivariate analyses have been used or because of only 9 sites, there would be too few degrees of freedom? What distribution are the p-values calculated on, or were they calculated via permutation tests?

As suggested, we have considered multivariate analyses and alternative combinations for the aggregation and combination of the rainfall metrics – and have long reflected on this commentary. However, and as recognised by the reviewer, the given dataset provided "too few degrees of freedom" to necessitate any such approach – hence p-values were

calculated on fitting data regressions and not permutation tests. Nevertheless, arithmetic addition of the rainfall metrics still resulted in important outcomes in our analysis. Albeit simple, these outcomes should not be discredited.

Line 10 onwards on NDVI. Need more information on how spatial results match the mine sites and how NDVI gives a good indication of plant early establishment in the mine sites. Although I may be missing something regarding how this has been worked out? Is the resolution good enough to only look at NDVI on the rehabilitated areas or does it include surrounding vegetation?
In the discussion changes to mine practice are mentioned e.g. irrigation, drainage etc that presumably aid vegetation establishment. If so, how can NDVI then aid in understanding the climate suitability unless the modifications were unsuccessful? More elucidation on why two one year El-Nino / La Nina periods were used and what months these actually corresponded to - note that on lines 26/27 pg 18553 these are referred to as one year El Nino periods whereas in this section they are referred to as lasting less than the calendar year. More justification is required for extending it to the year? My understanding is that El Nino's / La Nina phases often last longer than a year too? Exactly how long did they last for this analysis?

As described in the general commentary above, we have clarified the application of NDVI in this study. Most of the write up on the remote sensing methods used have been moved to the appendix. Additional text has been included that will provide clarification for your questions which include further explanations of: i) the relationship between NDVI pixel values and the mine site location and surrounding vegetation, ii) length of El Nino/La Nina periods chosen, and iii) the justification for choosing a single year.

RESULTS
On page 18554, line 25 it is noted that there are two non-significant rainfall criteria for predicting NDVI. What happens if these criteria are taken out of the overall suitability index? Does it change the conclusions? Note that the two non-significant criteria are also the two that could be most open to sensitivity analysis as mentioned earlier. Can you cite any literature for the inference that the non-significant results relate most to short term vegetation development. Again, this also relates to whether there are different criteria for establishment vs continued growth?
[...]
DISCUSSION
The discussion is on the whole structured well, and I thought the statements surrounding the importance of seasonal intensity in rainfall (page 18557, lines 12 onwards) were useful. In this vein, I think you could make a strong message of the Discussion that no one rainfall parameter allows you to predict suitability for rehabilitation but that an index is required. However, others may argue that one of your Figures shows annual rainfall correlates well with NDVI and, providing NDVI does indeed relate to early establishment potential (as per my earlier points), then is there any need for the index? Again, I think a stronger demonstration needs to be made of NDVI's potential to describe early establishment given statements such as page 18556 line 13 "particularly regarding the early-establishment of plants among post-disturbance ecosystems" Page 18558 line 17 - a southern hemisphere perspective. This may need emphasising or rephrasing to make it general. I find it surprising that this closing perspective in terms of how to deal with seasonal rainfall in mining rehabilitation has not been discussed before.

These are critical points. We closely revised these aspects throughout the Results and Discussion and outlined the further short-term rainfall/early establishment challenges with support from existing cited literature (Section 4.1, pg.10-11):

"At its extremes, the manner in which post-mining sites were deemed most-to-least sensitive across eastern Australia has provided a rather predictable depiction of how broad-scale rainfall patterns shape climate boundaries among arid central-inland vs. temperate coastal-hinterland locations. Evidently, regular rainfall and relatively short periods of water-deficit are common characteristics of favourable climate conditions, whereas prolonged seasonal drought with high variation, and frequently occurring intense rainfall events (storm or cyclone events) are primary characteristics of susceptibility; particularly regarding the early-establishment of plants among post-disturbance ecosystems. These fundamental relationships can be illustrated conceptually in our climate sensitivity matrix (Fig. 4) which identifies a range of climate scenarios (including moderately sensitive outcomes) in relation to the combined effects of differential rainfall availability and seasonal variation [...]"

As for issues pertaining to NDVI, we've clarified its application in the study (as in response to General Commentary, above) and thereby tightened the scope of our outcomes to avoid interpretive over-reach.

In Table 1, do the different primary commodities lead to fundamentally different landform elements and if so, does that have any influence on suitability for rehabilitation?

In brief, yes. The different commodities mostly lead to significant disturbance to the abiotic system and require extensive (and costly) landform reconstruction. Nevertheless, our study focuses on the overarching challenges associated with climate (somewhat) regardless of the specific rehabilitation measures in questions. By changing core terminology (i.e., from suitability to sensitivity) we've avoid confusion regarding these differences.

TECHNICAL POINTS
A number of long sentences and some use of jargon, particularly in the abstract. I think it would aid understanding if these were shortened and simplified to communicate the fundamental message.
Some typographical points: Page 18548, line 4 - should be "annual" rather than "annually" Ibid, Line 22 - weather-bound is unclear. Page 18552, line 4 and 5 - think should be "began

We have taken care to avoid long sentences and revise syntax in order to improve flow and readability.