Dear Prof. Wickert,

We greatly appreciate your time in reviewing our article. The comments are constructive and very helpful and improved the quality and hopefully the impact of the paper.

We agree that essential information was missing, especially regarding the genesis, and the resulting complexity, of the palaeovalleys in central Australia. This has been addressed in a new **Section 2.1 Genesis of palaeovalley systems in central Australia**. Next, we appreciate your comment about generalizability of our methodology, specifically about the lack of information that allows the reader to make their own determination about the potential usefulness of this methodology for their landscape/application. This has been addressed by yet another new **Section 4.3 Generalisation**. We trust that these major modifications have made the manuscript acceptable.

Other minor comments have been addressed as well. The detailed responses are listed below, with the changes marked in the annotated manuscript.

We look forward to discuss with you any further queries.

Thanks again.

Best Wishes,

Zhenjiao

Responses to editor

Generalizability. Your study site has been quite tectonically stable for a very long time, and was never glaciated. Therefore, the geomorphic evolution is straightforward: erosoin of bedrock uplands produces sediments that accumulate in basins, and geological and climatic processes lead to the formation and filling of valleys. In this case, these valleys have been loci of deposition throughout the Cenozoic. However, Reviewer 2's comment notes that your situation does not match that of many other parts of the world, where surface topography and subsurface structure are not linked. As an example that I know well, note Minnesota (USA)'s depth to bedrock (https://mngs-umn.opendata.arcgis.com/}, topography (http://arcgis.dnr.state.mn.us/maps/mntopo/), and aeromagnetic anomalies (https://mngs-umn.opendata.arcgis.com/app/the-aeromagnetic-database). You would not obtain the same answer with your methods here, where I am! However, as you note, it works for both your training and test regions, both of which are within the same physiographic and geological region. Therefore, I think that you need to couch your algorithm as an approach to find and characterize valley fill when topography is still present. And in so doing, it might be good to look into the work of Mey et al. (2015: {https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2014JF003270}, who also address this problem.

Reply: A new section **4.3 Generalisation** has been introduced to address the above comments (**Line 356-415**). We acknowledge that there was insufficient description of the site features for readers to assess to what degree the methodology could be transferred to other regions. We trust this has now been addressed, giving the reader guidance about topographical and other landscape features that allowed a successful application of proposed model using topographic predictors such as MrVBF.

4.3 Generalisation

The geomorphological evolution of palaeovalley systems is never straightforward. Our study site, for instance, remains tectonically active, although not in a manner where more recent tectonism leads to hills and ranges. Rather the neotectonism leads to changes in the hydraulic conductivity of aquifer systems. In some instances this results in marked changes in the conductivity structure across faults which transect palaeovalley systems (see, for example, Munday, 2020a; Munday et al., 2001). Furthermore, application of AEM for mapping buried valley systems has been successful in several other areas across Australia, each with their own evolutionary intricacies (Davis et al., 2016; Magee, 2009; Roach et al., 2014) with the key commonality being a very low topographic gradient. AEM has also been used in northern Europe, Canada and the US for similar purposes, albeit with different electrical conductivity structures. The application of AEM to map palaeovalley systems in many parts of the world has been successfully demonstrated.

Indeed, palaeovalleys occur beneath the glaciated landscapes of Northern Europe, Canada, and the Northern USA. When filled with coarse-grained permeable sediments, these valleys – as their Australian counterparts - represent potential sources of groundwater. In Northern Germany, shallow strata deposited during Quaternary times developed into palaeovalley systems characterized by a ground floor topography filled by fine grained marine and glacio-marine sediments. In these systems, AEM was successfully used to derive a detailed 3D geological model of the 350-m deep and 0.8-2 km wide valley infill (Siemon et al., 2006). Similar buried valleys with heterogeneous infill have been reported for Denmark, with typical dimensions of 0.5-4 km wide, and 25-350 m deep; their lengths varies from roughly 30 km for onshore structures to 100 km for offshore systems (Høyer et al., 2015; Jørgensen et al., 2003). In our study site, the burial depth of palaeovalley infills ranges between 5.0 to 250 m, with the typical width from 0.1 to 2 km.

In southern Manitoba, Canada, Oldenborger et al. (2013) used a combination of airborne time-domain electromagnetics, electrical resistivity and seismic reflection to map the complex buried valley morphology with nested scales of valleys at a level of detail sufficient for groundwater prospecting, modelling and management. Korus et al. (2017) demonstrated that AEM can be used effectively in environments like the glaciated Central Lowlands of Nebraska (USA) to identify sedimentary architectural units with a high degree of lithological heterogeneity. These systems were tens of meters deep and 100 m to more than 1000 m wide.

All these valley-infill systems are characterized by a multi-phase history of glaciation and buried valley genesis. The palaeovalley systems in our study area and the broader Musgrave Province/Great Victorian

Desert also have a multi-phase history, albeit with somewhat different processes across more extended timescales. However, the geologic/geomorphologic complexity of the Australian palaeovalley systems is therefore no less. Importantly, geophysical inversions across a wide range of palaeovalley systems have consistently delivered realistic geologic profiles, albeit defined by the geo-electrical properties of the fill materials and the water contained therein (Davis et al., 2016; Magee, 2009; Roach et al., 2014; Soerensen et al., 2016). They thus form a sound basis for subsequent developments such as deep-learning model for predictive purposes. Among the many potentially suitable geoscientific data sets for deep-learning-based prediction of palaeovalley boundaries and internal structure, topographic information was shown in this study to be a suitable predictor across a large test area (6400 km^2).

Valleys are, by definition, low points in the landscape and therefore topographic information is pivotal when mapping palaeodrainage patterns. In Australia, with its long-term tectonic stability, the topography of drainage systems has survived for very long periods of time. The presence of Mesozoic- Cenozoic pre-existing valleys has survived in the new landscape, because both erosion and deposition rates are extremely slow. These factors have combined to preserve many ancient Tertiary palaeodrainage patterns and in most instances palaeovalleys are still actual valleys, eventhough relief is subdued. Digital elevation models are very effective in recognising such Tertiary palaeovalleys and related features because the modern and Tertiary geomorphologies are usually related, both spatially and genetically (Magee, 2009).

Further characteristics of the palaeovalley landscape of the Musgrave Province are the extensive aeolian sandplains and sand dunes that overlay the valleys; groundwater calcrete and gypsum-rich playa sediments are evident in palaeovalleys where sand dunes are absent (Magee, 2009). These sand dunes were deposited around 200, 000 years ago (Krapf et al., 2019). The thickness of these sand deposits varies, but drillhole investigations revealed the boundary between overlying sandplains and palaeovalley to be around 30 m for major valleys to 10 m for tributary channels (Krapf et al., 2019). As a result, the palaeovalleys have only a subtle surface expression in today’s landscape. As we have demonstrated, detailed topographic data such as high-resolution MrVBF can be successfully used to detect such subdued surface expressions and infer the presence of buried systems.

In summary, palaeovalley relief is minimized and concealed by infill material, overlying sediments and the formation of playas (salt lakes). As a result, DEMs and its derivatives like MrVBF do not always permit the direct interpretation of palaeovalley boundaries, while the palaeochannel facies are even more difficult to infer (Hou et al., 2000). However, palaeodrainage systems in our study area mostly coincide with topographic lows characterized by MrVBF values between than 4 and 7 (inter-quartile range) (Fig. 4c).

So far, our deep-learning model has been tested and validated in the Great Victorian Desert only, noting the areas for training and validation were considerable in size, each 6400 km^2. Based on the characteristics of the palaeavalleys and the topographic features of the surrounding terrain discussed above, potentially suitable areas for further model testing can be identified. Note that the proposed model is not restricted to topographic input parameters only; any parameter that can be correlated with palaeovalley structure and features has potential to be used for predictive purposes. Therefore, the model developed here mainly serves as a generic framework that has applicability also in other areas, with input data not restricted to topographic information but also including remote sensing and geophysical data (Hou and Mauger, 2005).

In addition, the Mey et al. (2015) reference was cited in **lines 50-52**. For clarification, the trained neural network in this study is merely workable for palaeovalley prediction in central Australia, i.e. the underlying relationship in the trained model should be consistent to those regions for which prediction are made. For the application in other regions and other subsurface structures of interest, the model will need to be re-trained based on the specific input and output datasets. Alternatively, other predictors may be used, such as remote sensing data. In this regard **Figure 2** is also modified to make the input data more general.

Fitting geological reality or a geophysical inversion? I am concerned about your note regarding the valley structure being complex and discontinuous. This is inconsistent with your description of fluvial deposits, which should form a continuous sedimentary package (i.e., streams flow consistently downhill). There are two options. One is tectonics and localized deformation that reshaped the contact between the basement rock and the overlying fluvial deposits. The other option, which I find to be more compelling, is that the geophysical inversions may be imperfect. Although I am not as familiar with aeromagnetic data, I believe that inverse-distance effects would cause narrow sections of the paleovalley to be harder to detect and/or to seem to terminate at a shallower depth than wider portions of the valley. If I am correct, then your ML-based fit would be tuned to the details in a depth-and-wavelength-sensitive geophysical inversion rather than designed to match geological reality based on lithological data and interpretations.

Reply: We included a new section (**Section 2.1**) describing the genesis of the palaeovalleys in central Australia. This will illustrate their complexity, including the role of tectonics and other processes. We also argue that at least for the larger structures, the model fit is primarily using an accurate representation of geological reality, recognizing that by its nature the geophysical inversion will always be a simplification of true geology.

2.1 Genesis of palaeovalley systems in central Australia

The genesis of the central Australia palaeovalleys of the Musgrave Province (including the Great Victorian Desert and the APY Lands as our study area) covers about 60 Ma, and started as early as the Mid-Late Mesozoic to Early Palaeogene (about 65 Ma ago) with the latest palaeovalley infilling completed during the Early to Late Pliocene (about 2.5-5 Ma ago). The palaeovalley history involves a sequence of fluvial depositional periods interrupted by marine incursions, with climatic boundary conditions ranging from warm and humid (Late Miocene) to more aridic conditions (Late Pliocene to Early Pleistocene) (Krapf et al., 2019).

Valley incision was preceded by deep weathering of exposed basement rocks in the Mid to Late Mesozoic (Alley et al., 2009). While timing of the incision is debated, Hou et al. (2008) considered that the first infilling of the palaeovalleys with sandy fluvial deposits occurred through the Late Mesozoic – Early Palaeogene (about 65 Ma ago) and was focused along long lived, and still active (Pawley et al., 2014), structural discontinuities within the faulted Mesoproterozoic crystalline basement rocks (Figure 1). In the subsequent Late Miocene to Early Pliocene (about 40-13 Ma), characterized by a warm and humid climate, both freshwater and marine environment reversals occurred with marine sediments being deposited, transitioning to brackish and freshwater lakes (playas) occupying the valley floor. During the Late Miocene to Early Pliocene (about 10-3 Ma ago), evaporation of these sediments led to the deposition of a gypsum layer which was accompanied by intermittent fluvial deposition. The second and final fluvial deposition with quartz-rich sands occurred during the wetter Early to Late Pliocene (about 2.5-5 Ma ago). After this, the sedimentation continued into the Quaternary, with deposition of fluvial and colluvial sediments across the aridic landscape. During the Pliocene – Holocene (about 4 Ma ago to present), sand plains and sand sheets developed as a result of aeolian processes (Krapf et al., 2019; Munday, 2020b).

As a result of this long history of land-forming processes, the valley structures of our study area are complex, with varying width and geometry (Krapf et al., 2019; Munday, 2020b). Whilst fluvial systems at the coarse spatial scale are “continuous”, at a finer scale they may be discontinuous – shifting braided channel systems resulting in pinching out of fine or coarse scale sedimentary packages, etc.(Krapf et al., 2019). A not insignificant role in the creation of lateral discontinuities was played by Neotectonics resulting from the reactivation of basement structures, which in the context of the APY Lands (our study site), created discontinuities in both sedimentation and valley development, and important to groundwater systems, formed hydraulic barriers in the overlying sediments. Such variations in width and depth of palaeovalleys can cause discontinuities when airborne electromagnetics are geophysically inverted, particularly if the valleys are smaller than the footprint (resolution) of the airborne system. However, most prominent are discontinuities in the lateral continuation of the conductivity features associated with the valley fill, particularly where major fault systems cross cut the primary orientation of the valley systems. These become particularly apparent at depth (>50 m) in the subsurface (see Munday, 2020b). This is attributed to the effects of active tectonics during the valley fill events.

The influence of neotectonism on the observed conductivity structure associated with palaeovalley fill sequences has been discussed elsewhere by Munday et al.(2001), while Munday et al. (2016) highlighted the role neotectonics may have played in influencing the patterns of the observed electrical conductivity structure in the Musgrave province of South Australia. These studies demonstrated the role faults, interpreted in the regional magnetics, play in influencing the presence of abrupt discontinuities in the modelled conductivity structure.

More important for the success of AEM in deriving palaeovalley features is the variation in the petrophysical properties of the valley fill materials. If those properties vary, then one can expect to see an airborne system varying in its capability to map continuity. The critical factor for deep learning (DL) applications is understanding what the DL-based fit is actually working with. That would determine whether one is fitting a geophysical expression of a geological system, which by its nature will be a simplification of true geology, or geological reality. The geophysical expression is well matched with the geological reality, albeit a simplification of geological reality determined by the resolution of the airborne system, and the geo-electrical properties of the target materials. For the larger palaeovalleys, where the conductive structures identified correlate well with alluvial fill of the valley systems, the geophysical data maps geology well. Consequently it is reasonable to argue that DL is fitting geological reality, but at the finer scales a mismatch may occur between geological reality and its geophysical expression. For both scales, however, the DL application will be affected by the underlying geophysical expression of the geology and the inversion approach used. In the latter case we employ a 1D Layered Earth Inversion (LEI) routine with lateral constraints. 1D assumptions in the inversion include the assumption that the earth consists of uniform, laterally extensive layers. At the scale of AEM system footprint and mapping scale of this study this assumption holds. Similar inversion assumptions have been successfully employed in other studies of palaeovalley systems using 1D inversion codes (e.g. Høyer et al., 2015; Korus et al., 2017). Davis et al. (2016) and Roach et al. (2014) reported on the successful application of 1D inversion approaches with AEM data for delimiting palaeovalley systems in Australian settings.

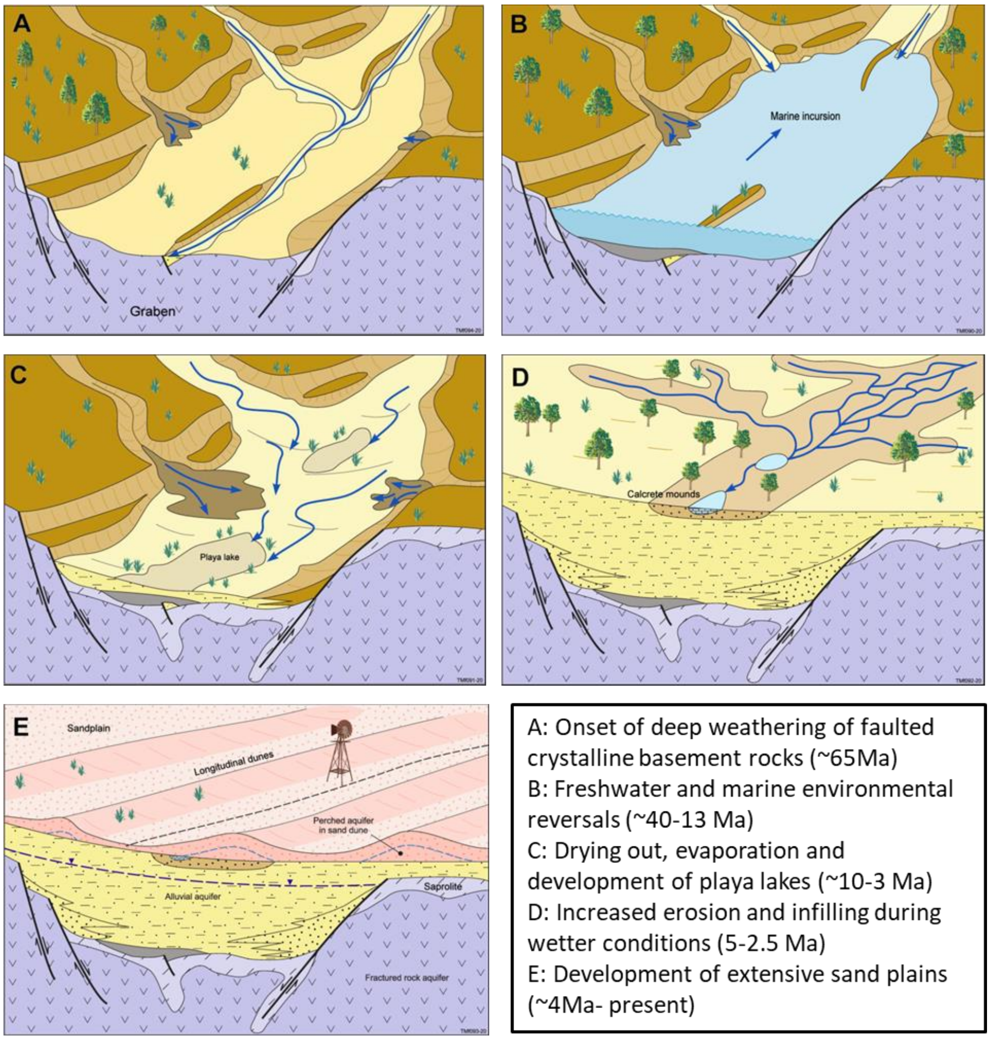

Figure 1. Conceptualised genesis of the palaeovalley landscape in the Musgrave Province in South Australia (after Krapf et al., 2019; Munday, 2020b).

We further made the following changes to support the description of the palaeovalley structure:

(1) **Figure 2b** is redraw to better describe the general lithofacies in the palaeovalley in this study area. The lithofacies feature coarse sands deposits gradually evolving to the clay-dominant depositions, overlying by the aeolian sands and silts. This represents the depositional environment changing from wet

to dry periods. The sudden lithofacies change in the previous image was quite misleading.

(2) It was found that the high bulk electrical conductivity values (EC) are a proxy for palaeovalley presence in contrast to the low EC of the bedrock; the higher EC, the higher probability of the palaeovalley presence (**lines 254-257**). We thus use an AEM-derived index to indicate the occurrence of palaeovalley. We also clarify that the model fit is based on the geophysical inversion rather than the lithological data (**lines 261-264)**.

Line 12. tomography → topography?

Reply: This is now modified in **line 13**.

Line 75. their → its

Reply: This is now corrected in **line 20.**

Line 79. "sandplain" is not a genetic term. Later in the paper, it seems that you indicate this to be aeolian. You should make this clear here, because it importantly indicates that you are inverting across a buried and preserved paleolandscape and its deposits. (You can expand upon this to in regards to the geological setting.)

Reply: The ‘sandplain’ is now replaced by ‘aeolian’ for clarification.

Line 82. “As is evident from an analysis of AEM images, the palaeovalley system has a highly irregular geometry with spatially varying depths to basement, and with heterogeneous infill resulting in lithologically controlled palaeovalley aquifers”. This is the core of my second major point, above.

Reply: This has been addressed in the **new Section 2.1**.

Line 85-87. “Our goal is therefore to develop an efficient and generic machine learning tool to express the relationship between an easy-to-obtain dataset and a more costly dataset for the specific purpose of detecting palaeovalley features that would facilitate the discovery of new groundwater resources in arid and semi-arid regions.” See the first point; I think that this might not be as generally applicable as you state, especially since you test and validate it in the same stable geological setting.

Reply: We agree that this statement was somewhat optimistic in regard to the success of DEM data as predictor variable, while the methodology itself is still sufficiently generic such that other data types can be used for prediction purposes. Text has been rephrased to emphasize that we are using this study area as an example to demonstrate the effectiveness of the method. For the application in other areas with other subsurface structures of interest, the model will need to trained again based on the problem and data at hand (**e.g. Line 56-59, Line 69-71 and Figure 2**). Also, while other predictor data may be more suitable than DEM data, the method itself is sufficiently generic that it can be easily adopted.

Line 90-91. “Such methodology is premised on the existence of a mechanistic connection between landsurface features and subsurface distribution of palaeovaleys”. This indeed is THE key limitation, and I think it deserves to be more clearly stated, as this will help readers at once recognize whether your approach is one that they could use in their system or not.

Reply: Thank you, good point indeed. The new **Section 4.3 Generalizability** addresses this in detail.

Line 103, Fig. 1, Is your use of "image" standard? Because to me, these are “arrays" or “layers of inversions", but not actually images.

Reply: Following the terminology in deep learning, we use the ‘image’ to express the input and output data. Considering this comment, the ‘3D array’ is annotated following the first appearance of the 3D image in **Figure 2 and line 184**, for clarification.

Line 154. This is your first mention of a CSIRO data set. What is it?

Reply: This is now explained in **line 253** and the link is given in the **code/data availability section.**

Line 157. I assume that this is the ``sandplain", per my above comment.

Reply: This is now replaced or clarified with ‘aeolian’.

Line 159. The MODERN valley bottom. In addition, you have not yet introduced the valley-bottom flatness index. It seems that this paragraph may be out of place, and should follow the paragraph below in order to provide the requisite information first. **Line 169.** Define how the MrVBF is calculated and the details of the source DEM (which data source, resolution, etc.). **Line 176.** MrVBF is a derived data product, not a data source in itself.

Reply: This part is now rewritten in **lines 218-254.**

Line 183. Not every deposit filling a valley is an aquifer. Is this formula sensitive to the nature of the sedimentary fill? Associated with that, the cross section in Fig. 2B indicates an aquitard between two aquifers.

Reply: Yes, the previous figure 2b is quite misleading, we now redraw **Figure 2b** to better express the general lithofacies in the palaeovalley.

Line 228. What is a fully connected layer in the encoder? Could you provide a bit more ML background to help the reader to understand why it is important and creating the observed result?

Reply: This is now explained in **lines 307-308.**

Line 287-289. I appreciate your desire to contribute, but do not think that your tool has generic applicability based on the points that I have raised.

Reply: This sentence is now rephrased in **lines 426-428.** We trust that the major changes made throughout the manuscript will make clear what landscape features are required for this DEM-based method to be have more general applicability.