**Ecological niche modelling of *Solanum tuberosum* L. under present and future climatic conditions**

***Methods in Biodiversity Analysis report - Jelle Dercksen (1561790)***

**Introduction**

*Solanum tuberosum* (L.), which is commonly known as the potato, is a widely cultivated plant species. The potato ranks as the fourth most important food crop globally. *Solanum tuberosum* originated from the Peruvian, Bolivian and Chilean Andes of South America and was introduced in Europe in the 16th century after which it spread around the world. Currently it appears in Africa, Asia, Europe, Oceania and North and South America. In few countries *S. tuberosum* is declared a noxious weed since every part of the plant, except for the tubers, are poisonous for consumption. GBIF lacks occurrence data of *S. tuberosum* in the central states of North America, eastern regions of Asia and southern regions of Africa (figure 1). The occurrence data in Europe largely consists of cultivated potatoes.

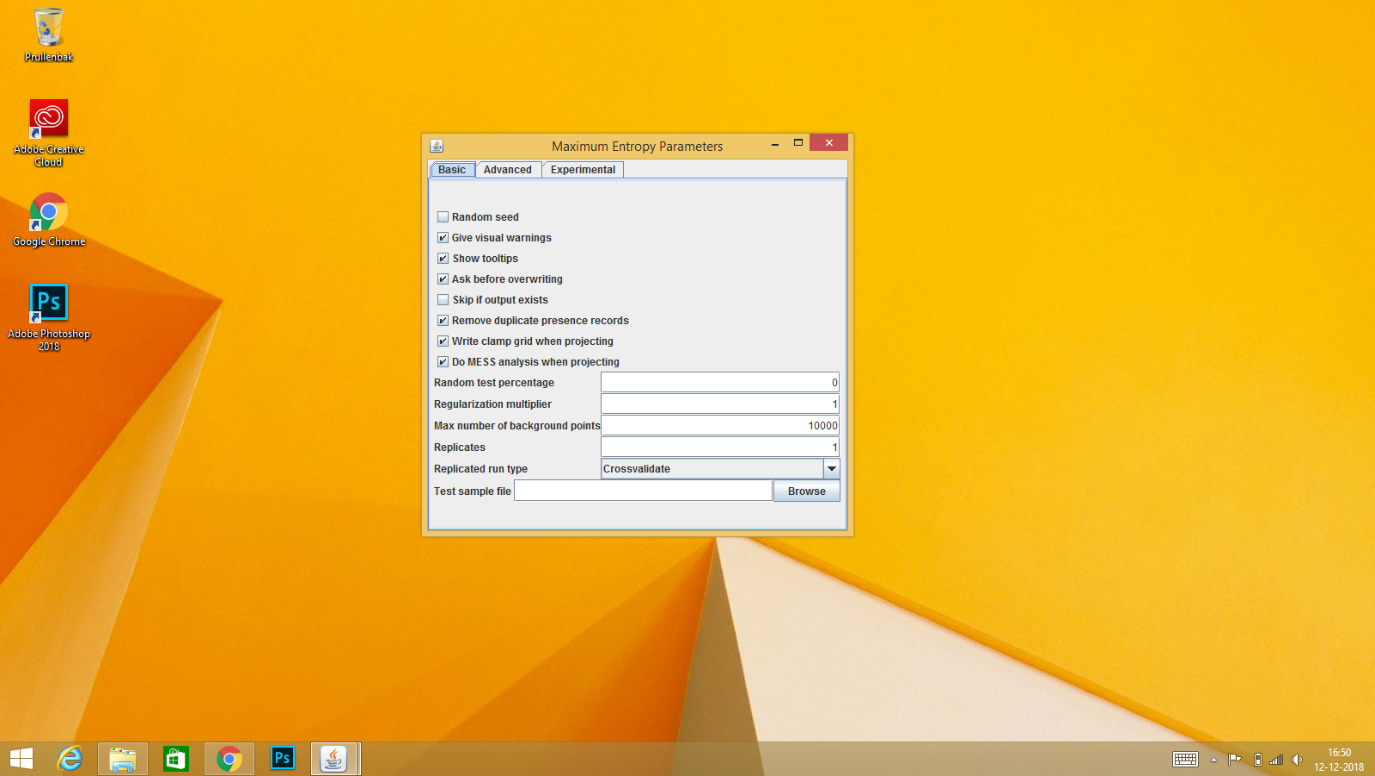
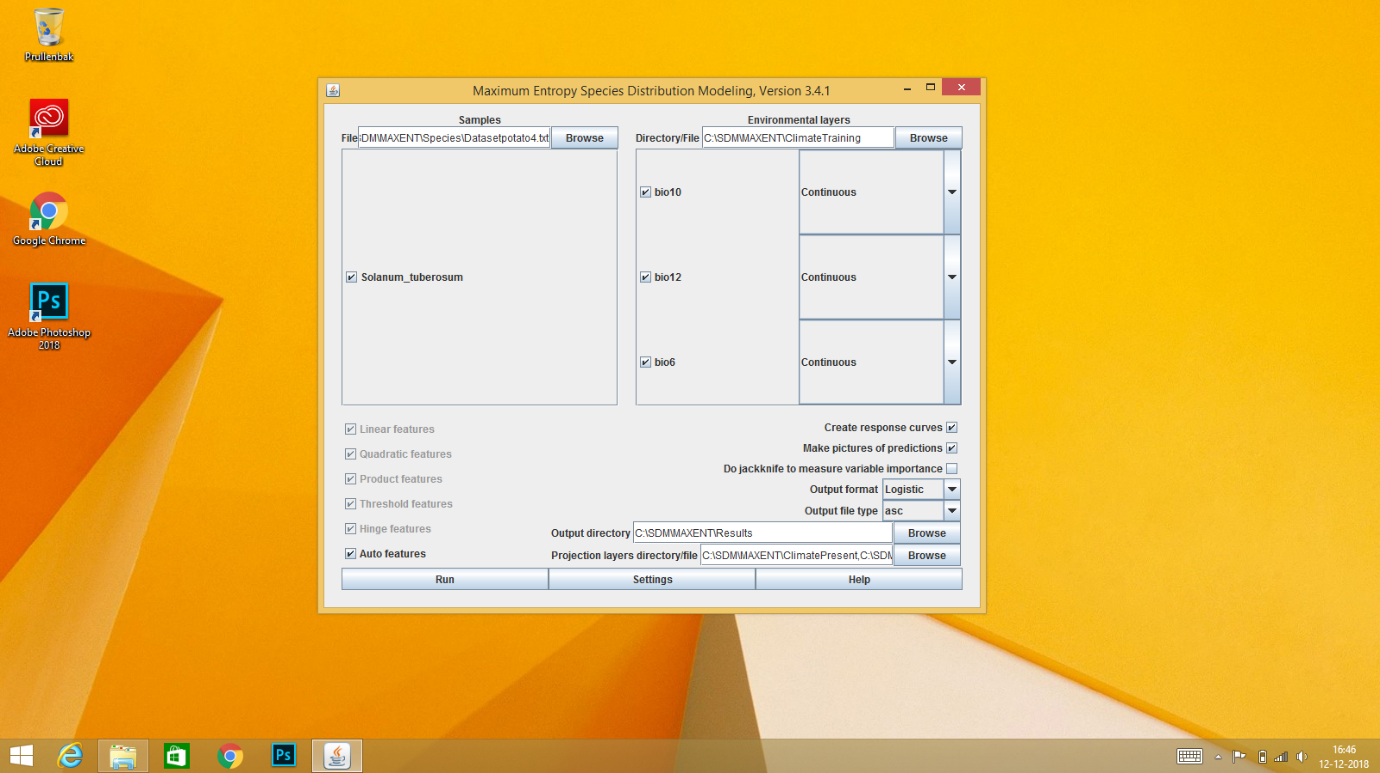


**Figure 1:** Present distribution of Solanum tuberosum (L) according to georeferenced GBIF occurrence data.

**Methodology**

Data from GBIF (see download references) was obtained after which MaxEnt (version 3.4.1) built the species distribution models for 2050 in the RCP 4.5 scenario with the use of bioclimatic variables obtained from worldclim.org at a resolution of 5 minutes from the HadGEM2-AO model. RStudio (version 1.1.463) was utilized to select the relevant variables, and clip the rasters to their bounding boxes by application of the Rscript provided on GitHub. The first dataset from GBIF consisted of 15,193 georeferenced occurrence records and the second extracted set was clipped on the range of Bolivia and Peru consisting of 2,849 georeferenced records. Multiple combinations of variables have been tested to keep the variables uncorrelated. The proposed relevant variables were Bio6 (min. temperature of coldest month), Bio10 (mean temperature of warmest quarter), Bio11 (mean temperature of coldest quarter), Bio12 (annual precipitation), Bio15 (precipitation seasonality), Bio19 (precipitation of coldest quarter). *Solanum tuberosum* is very sensitive to heavy frost (hence bio6, bio11) but thrives in cool conditions where tuber formation halts around 27°C (bio10). Moreover, *S. tuberosum* requires uniform irrigation throughout the year for development (bio12, bio15) where water also aids in cooling the soil. The seed pieces of potatoes in cold, damp soils are likely to rot before being able to grow however (bio19). In the first run a combination of variables bio6, bio10, bio12 was tested. During the second run, a combination of variables bio11, bio15, bio19 was attempted. In the third run bio6, bio10, bio12, bio15 and bio19 were combined. Each of these compositions had comparable AUCs (0.802, 0.803 and 0.807 respectively), where AUC values higher than 0.8 are considered good. A fourth run was performed afterwards utilizing data from *S. tuberosum*’s native range (Peruvian-Bolivian Andes) utilizing the variables bio6, bio10 and bio15 observing an AUC of 0.909. The native range is suitable for research since it contains the majority of genetic diversity. It is meaningful to maintain the gene pool for e.g. the crossbreeding towards stronger cultivars.

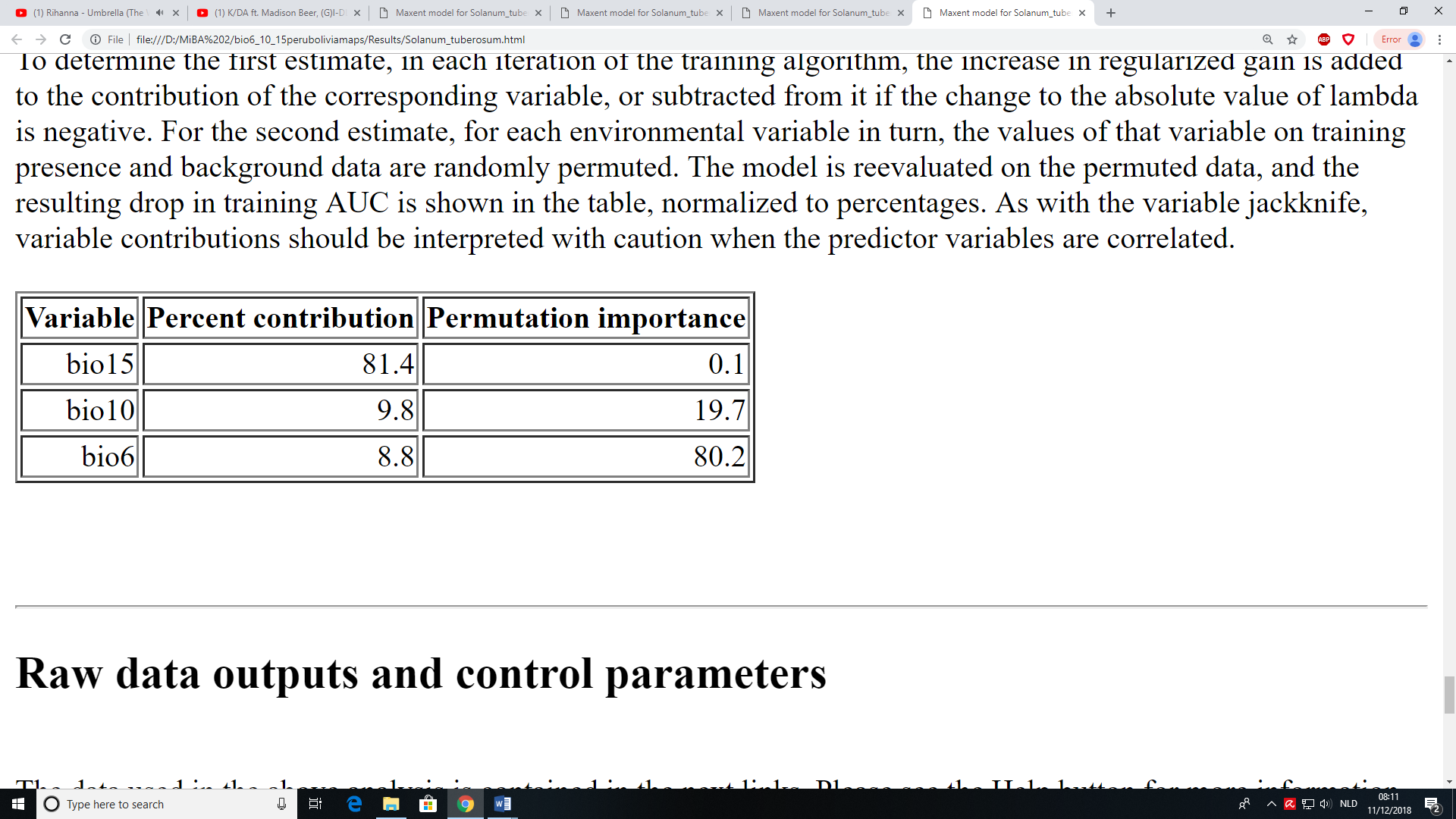
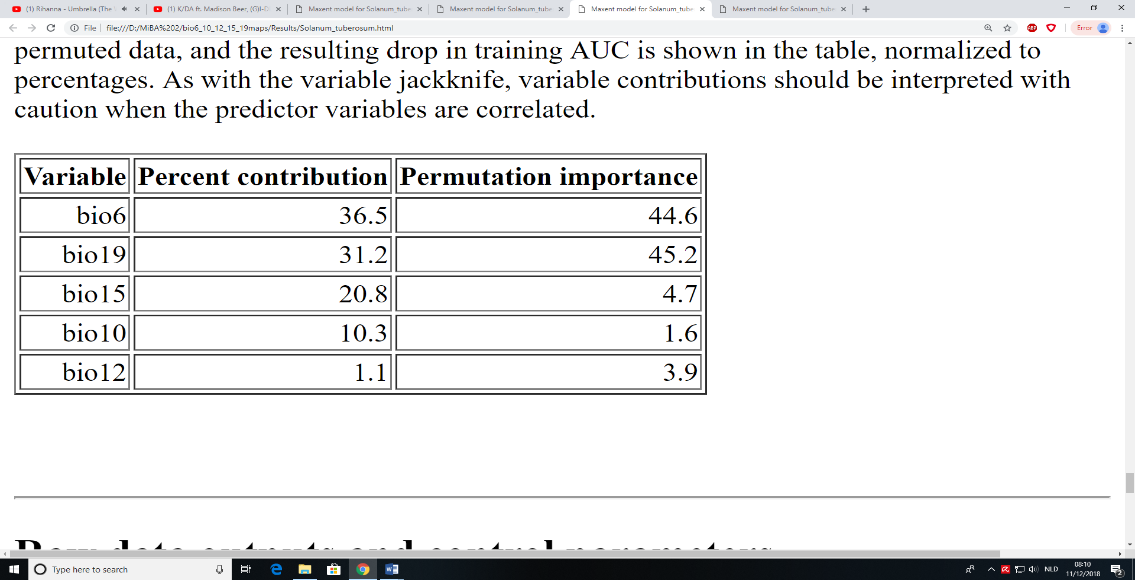
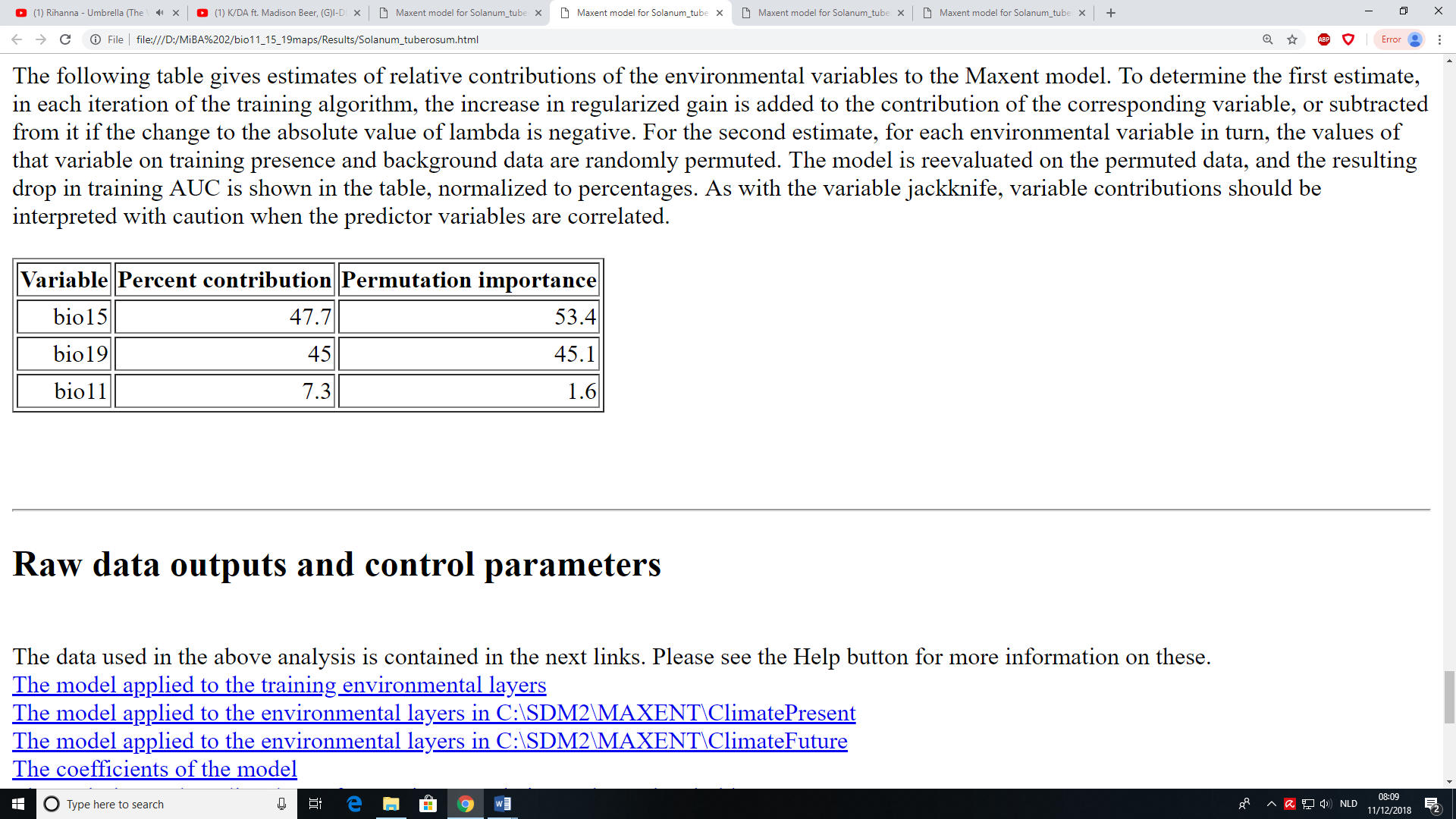
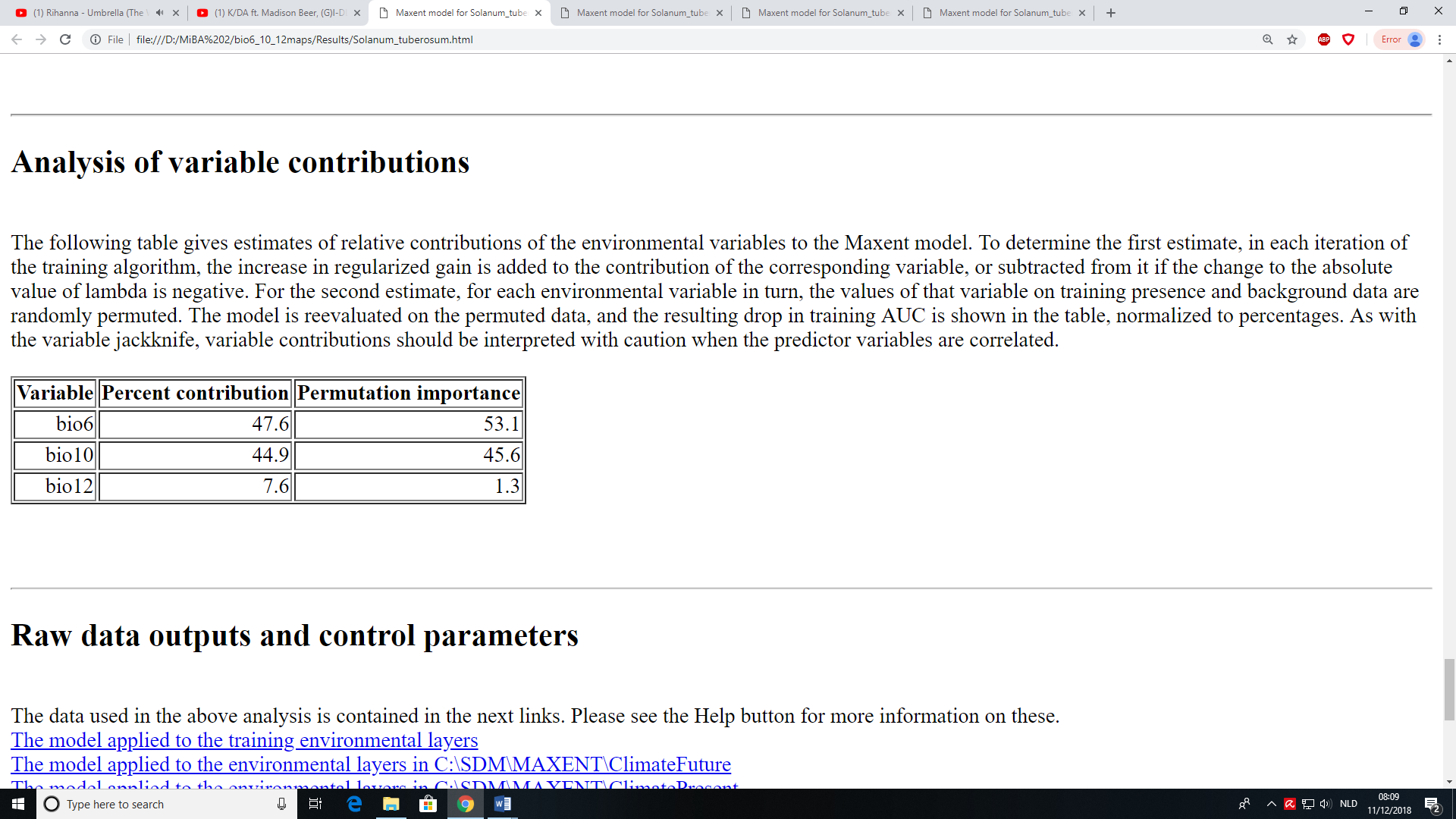
Identical settings were used in the MaxEnt application for these four individual combinations of variables, creating response curves and the output format being logistic. The settings are to be found in figure 2.



**Figure 2:** MaxEnt settings utilized in all runs

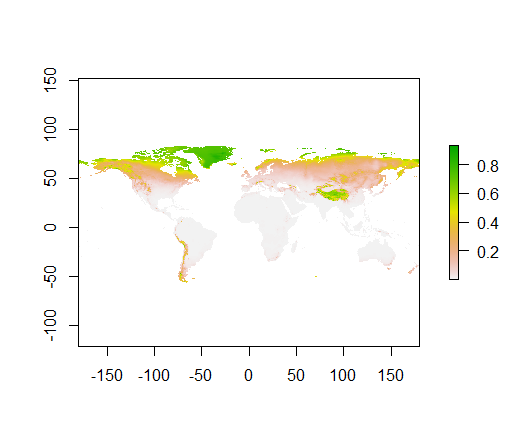
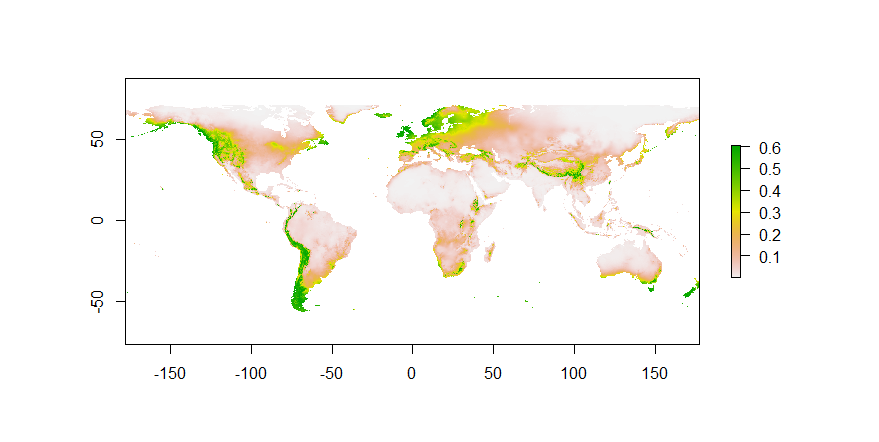
**Model output**

The SDMs constructed by MaxEnt were visualised, creating i.a. the following maps (figure 4, 5). Additionally ROC curves were constructed with the their respective aforementioned AUCs (figure 6). All variable importance tables are included in figure 3. Since multiple runs of variables were performed, not every visualised SDM-iteration could be included. The first three attempts utilized the same dataset and yielded comparable results. The fourth run (clipped to Peruvian-Bolivian Andes) yielded completely different results as can be seen in the differences between 4a, 5a compared to 4b, 5b.



**Figure 3:** Variable importance table of four individual variable compositions. Top-left: run 1. Top-right: run 2. Bottom-left: run 3. Bottom-right: run 4 (Peruvian-Bolivian dataset).

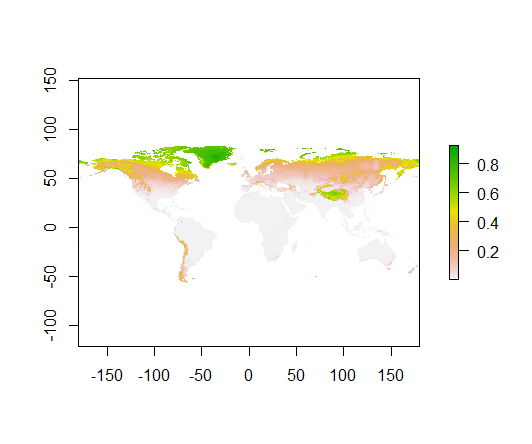
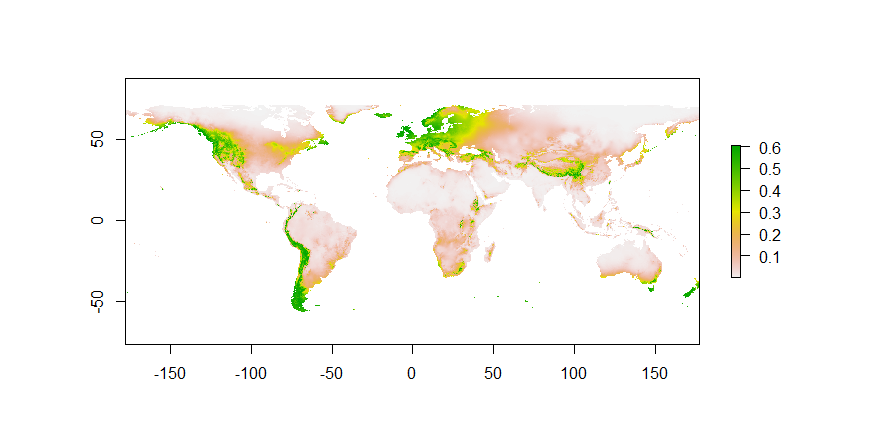
**Figure 5:** Future occurrence chance map of Solanum tuberosum. (A) Run 1 which has similar results as runs 2 and 3. (B) Future distribution is projected onto the arctic circle (run 4).



**B**

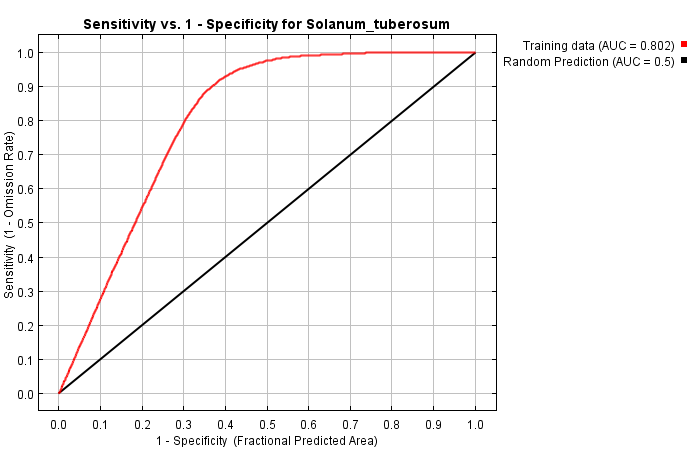
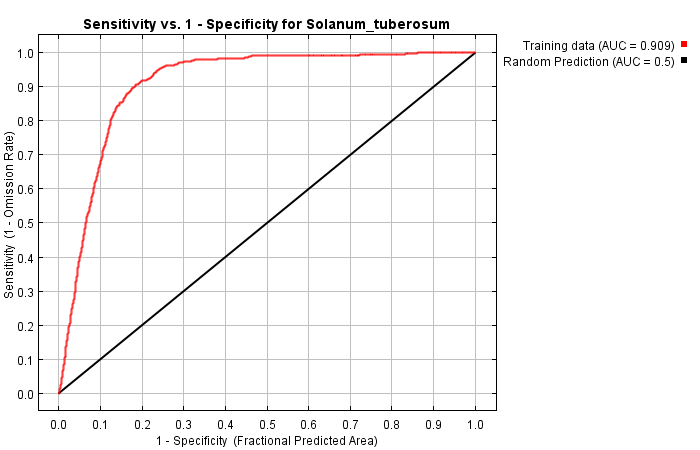
**A**

**Figure 4:** Present chance of occurrence map of Solanum tuberosum. (A) Run 1 which has similar results as runs 2 and 3. (B) Current distribution is projected onto the arctic circle (run 4).



**B**

**A**



**Figure 6:** Receiver operating characteristic (ROC) curves for (A) run 1 and (B) run 4 with their respective AUCs.

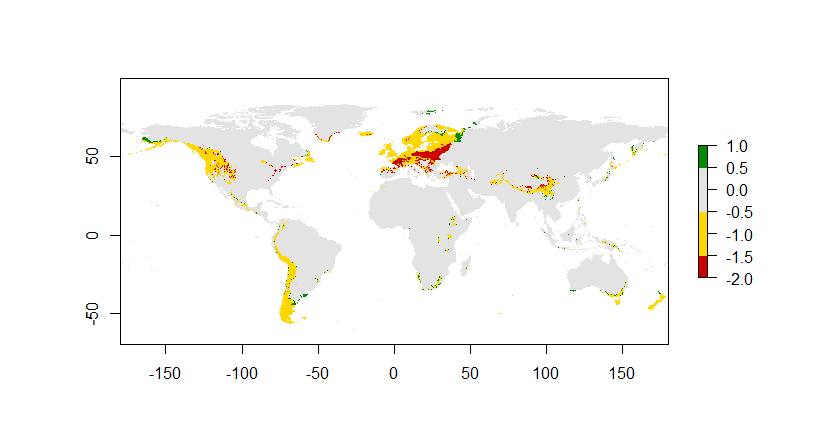
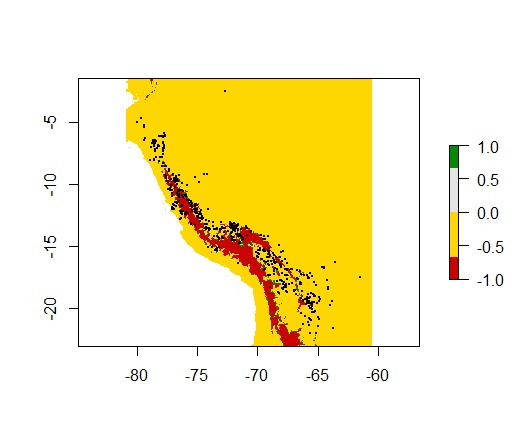
**A**

**B**

**Response to future scenario**

The relative change of habitat suitability as a result of the selected bioclimatic variables were mapped (figure 7). In this map we can see comparable change as between figures 4 and 5.

**Figure 7:** Future distribution change maps. (A) Habitat unsuitability arising throughout southern and eastern Europe (run 1 with similar results as runs 2, 3). (B) Illustrating S. tuberosum-habitat becoming unsuitable in the Andes (run 4).



**A**

**B**

**Biological interpretation**

According to the SDMs, south-eastern Europe and the Andes seem to become unsuitable for *S. tuberosum.* However, there are a few things to take into account when interpreting these SDMs. There is no distinction between different wild subspecies or genetically diverse cultivars. One cultivar could react better to climate change than another. Hence the effect of climate change by these variables can not be observed on a potentially wide range of genetic diversity. The geographic distribution provided by GBIF is dubious, as mentioned in the introduction, because large amounts of data are missing. Data on soil type and for instance salinity is also excluded whilst it is essential for a suitable habitat. In addition, the cultivated potato plants are intentionally grown under their preferred conditions. This means that climate variables such as precipitation do not represent the amount of water available for *S. tuberosum* through irrigation. The conditions of these agricultural regions however are extrapolated as favourable since the potato occurs there. This creates a misleading preference toward regions representing the same conditions as those of agricultural areas for future and current potato occurrence.

Moreover, during construction of a SDM the realized niche of a species is utilized as current distribution, whereas the future distribution is then based solely on abiotic influences neglecting biotic interactions and dispersal limitations. For a more accurate projection, dispersal and biotic interactions need to be added into the SDM. Lastly, only presence data is utilized, making the AUC a questionable validator of the SDM. This issue could be fixed by testing the SDM against a null-model.

Even though the AUC of run 4 (figures 4b, 5b, 6b and 7b) is high (0.909), the model is not representative of reality. Figures 4b and 5b represent the simulated chance of occurrence of *S. tuberosum*, however it is not reasonable to believe that *S. tuberosum* could flourish in the arctic circle since it does not withstand frost at temperatures under -3°C. In this model bio6 (min temperature of coldest month) was taken into account, though since it was trained with data from high altitudes this variable was not recognized as a limiting factor (see 8.8% contribution in run 4, figure 3). Taking all this into account, since *S. tuberosum* is sensitive to higher temperatures, climate change will most likely have an impact on the suitable range in its natural habitat (the Andes). It will most likely disperse to slightly higher altitudes thereby compensating for the slowly rising temperature. In figure 7a the southern parts of Europe *S. tuberosum* will become unsuitable, according to SDMs of runs 1-3, after which it might disperse north to seek refuge in cooler regions. Most of the potatoes occurring in these regions are cultivated. Hence, a lot of the future “unfavourable” conditions will most likely be compensated in these regions.

Not only is habitat suitability involved in the distribution of a species, it is also a matter of humans and terrain allowing a species to traverse area in a relatively small timeframe. On that note, human-caused spread of a species happens at an extremely quick pace. So if a species needed to traverse thousands of kilometres toward a suitable area in these next 32 years it would need to be by hand of mankind instead of by natural dispersal. I do not think these SDMs are useful in detail since a lot needs to be added/taken into account, but I do think that relevant trends are recognizable.

**References**

GBIF.org (05 December 2018) GBIF Occurrence Download <https://doi.org/10.15468/dl.olw5hw> (all georeferenced occurrence data)

GBIF.org (09 December 2018) GBIF Occurrence Download <https://doi.org/10.15468/dl.3jtqhw> (georeferenced Peruvian-Bolivian Andes occurrence data)

Solanum tuberosum (potato) datasheet. (2018, September 27). Retrieved December 9, 2018, from <https://www.cabi.org/isc/datasheet/50561>

Government of Canada,Canadian Food Inspection Agency, & Plant Health and Biosecurity Directorate. (2015, May 24). The Biology of Solanum tuberosum (L.) (Potatoes). Retrieved December 9, 2018, from <http://www.inspection.gc.ca/plants/plants-with-novel-traits/applicants/directive-94-08/biology-documents/solanum-tuberosum-l-/eng/1330982063974/1330982145930>