Kotze JDF, Beukes HD, Seifert T (2019).
Essential environmental variables to include in a stratified sampling design for a national-level invasive alien tree survey
iForest – Biogeosciences and Forestry – doi: 10.3832/ifor2767-012

Supplementary Material

Appendix 1 – Description on the approach used to filter IAP species within the SAPIA database.

Only species records that occurred within South Africa were included in the species selection process, resulting in 550 unique species. IAP species occurring within these areas have managed to establish, survive, reproduce and spread unassisted resulting in them remaining in the database. Where the proportion of total occurrence records for a specific species classified as belonging to the disturbed class, thereby not occurring in natural/semi-natural areas, was more than 75%, that species was excluded from this study. Only species that were associated with more than 20 QDS blocks were deemed as having a significant spatial range to contain meaningful environmental and species abundance gradients. To allow for the effective association between IAP species and especially soil environmental variables, only woody species and therefore perennial species were included. Annuals were excluded for their often strong and opportunistic association with favourable seasonal climatic conditions such as good rainfall rather than for instance soil variables. All waterweeds were excluded because their distribution is determined by the availability of waterbodies rather than any other environmental variables. Succulent species such as those belonging to the genera *Opuntia* were excluded due to their water storage ability. Species occurring within riparian zones were classified as riparian species, whilst those occurring within the remainder of the landscape outside of riparian zones were classed as landscape species. If a species' total proportional landscape position was less than 75% for both the riparian and landscape categories across all affected QDS blocks, it was considered as not being preferential to any of the two landscape types and remained in the database (Nel et al. 2004). Consequently, the qualitative abundance classes associated with each species record per QDS were assigned numeric class values that were provided by Henderson (2007) and which represented the approximate actual number of plants observed per 10 km transect, namely "Rare" was equivalent to 1 plant, "Occasional" to 10 plants, "Frequent" to 50 plants, "Abundant" to 200 plants and "Very abundant" to 1 000 plants. This allowed the frequency of occurrence of a unique species per QDS block to be combined with its associated abundance values and to be summed per QDS for that species. Species records with a combined abundance value per QDS block equal to or more than the class "Frequent" remained in the database. The largest proportion of the remaining species overlapped geographically by being recorded within the same QDS blocks, resulting in species being spatially mutually exclusive from each other for only a limited amount of QDS blocks. Therefore different species within the same QDS is associated with the same environmental variables for that QDS. In other words, different IAP species are equivalent to one species in terms of association with their environment within a specific QDS block alone. This overlap was addressed by comparing the spatial distribution of the remaining species with each other by determining the proportion of QDS blocks occupied by a species in relationship to the total number of remaining QDS blocks for all species. Species with a proportional coverage of 20% or more and that were 50% or more geographically mutually exclusive from each other were selected. The combination of maximum spatial distribution at a species level with minimum overlap at a QDS block level led to the selection of three species, namely *Acacia cyclops*, *Acacia mearnsii* and *Prosopis glandulosa*.

Appendix 2 – The stratification procedures followed of environmental variables.

Each of the environmental variables, except for terrain morphological units, were reclassified into three classes to obtain a gradient ranging from low to medium and finally high. This stratification was based on an equal area classification to ensure an unbiased equal probability design at a main effect level. Terrain morphological units were reclassified into three classes, namely the lower lying valley bottoms, middle areas as footslope and the higher lying areas in the landscape as a combined midslope/scarp/crest unit. The two aggregation units or spatial scales at which this stratification was done, was the complete study area, South Africa, and the tertiary catchment delineation, 274 in total. The reclassified environmental variables were geographically intersected in a cumulative stepwise manner by increasing the number of variables involved one at a time until all 14 variables were spatially intersected for the two aggregation levels.

Fig. S1 - Approach followed to filter the SAPIA database. The numbers in brackets are the remaining number of species after each filter step.