Predicted suitable sites for NZ sea lion breeding colonies on mainland New Zealand: table of data fields

Names and descriptions of data fields within the NZ sea lion suitable sites map, and how they apply to management.

Data field name	Description	Management application
Site identification		
id	Numerical ID, corresponding to the row number.	To differentiate among sites; site queries.
site_ID (site identification number)	Three-letter + number-coded ID for each site.	See **id**.
DOC_region	DOC operations region where a majority of the site is located.	To query and locate sites by management region.
region	Name of the NZ region where the site is located.	See **DOC_region**.
main_isld (main island)	North or South Island (includes Stewart Island).	See **DOC_region**.
X	X centrepoint coordinates (UTM).	To map sites.
Y	Y centrepoint coordinates (UTM).	See **X**.
Size		
area	Area of the site in km^2.	To prioritise or evaluate sites by size or capacity.
S1_area_pc (S1 suitable area coverage (%))	Proportion of a site (in percent) that is environmentally-suitable for the first behavioural state in the breeding season, S1 (breeding), based on the multi-state SDM.	To assess the availability of a habitat for each behavioural state.
S2_area_pc (S2 suitable area coverage (%))	Proportion of a site (in percent) that is environmentally-suitable for the second behavioural state in the breeding season, S2 (transition), based on the multi-state SDM.	See **S1_area_pc**.
S3_area_pc (S3 suitable area coverage (%))	Proportion of a site (in percent) that is environmentally-suitable for the third behavioural state in the breeding season, S3 (dispersion), based on the multi-state SDM.	See **S1_area_pc**.
Model uncertainty		
MESS_class (multivariate environmental similarity surface grid mean value class)	Classified mean multivariate environmental similarity surface grid value. A measure of extrapolation impact on an area's prediction, based on similarity comparisons with the SDM training area (Auckland Islands). Five classes describe extrapolation levels as follows: 1) none (values 0 - 100) 2) low (values -100 - 0) 3) intermediate (values -500 - -100) 4) strong (values -1000 - -500) 5) very strong (values -1600 - -1000)	To critically review predicted sites; sites with MESS class of 'none' to 'low' have qualities that are most similar to the Auckland Islands; sites with 'intermediate' to 'very strong' MESS classes have qualities that are dissimilar to the Auckland Islands, so there could be errors in the predictions.
MOD_md (mode of the most dissimilar variable)	Mode (most common) of the most dissimilar variable for a site, corresponding to the MESS grid.	To determine the variable that is the most dissimilar from the training area (Auckland Islands), which caused extrapolation.

S1_uncrt (uncertainty for the S1 prediction (%))	Coefficient of variation (CV) value (in percent) for the S1 prediction, estimating how far a prediction at a site deviates from the average prediction in the model-building process.	To gauge uncertainty in the SDM's prediction of habitat suitability. Relatively low percentages indicate low uncertainty (more reliable predictions), and relatively high percentages indicate high uncertainty (less reliable predictions).
S2_uncrt (uncertainty for the S2 prediction (%))	Coefficient of variation (CV) value (in percent) for the S2 prediction, estimating how far a prediction at a site deviates from the average prediction in the model-building process.	See **S1_uncrt**.
S3_uncrt (uncertainty for the S3 prediction (%))	Coefficient of variation (CV) value (in percent) for the S3 prediction, estimating how far a prediction at a site deviates from the average prediction in the model-building process.	See **S1_uncrt**.

Restoration features		
S1_limit (mode of limiting factor for the S1 prediction)	Mode (most common) of limiting factor (or, model variable) for the S1 prediction.	To evaluate the variable that limits the suitability of a site for that behavioural state; if this variable's values are improved, then site suitability can improve.
S2_limit (mode of limiting factor for the S2 prediction)	Mode (most common) of limiting factor (or, model variable) for the S2 prediction.	See **S1_limit**.
S3_limit (mode of limiting factor for the S3 prediction)	Mode (most common) of limiting factor (or, model variable) for the S3 prediction.	See **S1_limit**.

Human impacts		
hum_im_pc (potential human impacts coverage (%))	Proportion (in percent) of a site that has potential human impacts, based on the multi-criteria decision analysis of 3D distances from sealed/unsealed roads and residential areas.	To examine the degree or potential NZSL interactions with humans by roads and residential areas; to prioritise areas for community engagement and outreach.
rd_sl_mi (minimum sealed roads distance (3D; km))	Minimum 3D path distance (in km) of a site from sealed (paved) roads.	To expand on the information from the multi-criteria decision analysis; vehicle collisions are a threat to NZSLs.
rd_unsl_mi (minimum unsealed roads distance (3D; km))	Minimum 3D path distance (in km) of a site from unsealed (unpaved) roads.	See **rd_sl_mi**.
fences	The presence or absence of fences within a site.	To assess barriers that could not be included in the SDM; presence of fences implies that there is less suitable area available for the NZSL than predicted.
graze_pc (grazing grasslands (%))	Proportion (in percent) of a site that has high-/low-producing grasslands for dairy and non-dairy grazing.	See **fences**.

Additional suitability		
in_watr_mi (minimum inland water distance (km))	Minimum Euclidean (straight-line) distance of a site from inland water bodies (lakes, ponds, streams; excludes inlets).	To ground-truth the predicted sites; this SDM variable is limited because it is lacking inlets, which are also important for thermoregulation.

in_watr_me (mean inland water distance (km))	Mean Euclidean (straight-line) distance of a site from inland water bodies (lakes, ponds, streams; excludes inlets).	See *in_watr_mi*.
in_watr_mx (maximum inland water distance (km))	Maximum Euclidean (straight-line) distance of a site from inland water bodies (lakes, ponds, streams; excludes inlets).	See *in_watr_mi*.
pine_pc (pine forest (%))	Proportion (in percent) of a site that contains planted pine (*Pinus radiata*) forest (as of 2016).	To account for additional, non-native forest types that are preferred by recolonising NZSLs on the mainland.
Locations of interest		
curr_NZSL (current NZ sea lion sites)	Names of known current (1990 to 2019) sites where females and/or pups have been sighted during the breeding season (December to March) that are within a 10km Euclidean (straight line) distance from a site.	To investigate or ground-truth areas near where current management actions are taking place; optional suitable locations where NZSLs are likely to be found due to proximity to existing pupping sites.
histr_NZSL (historic NZ sea lion sites)	Names of known historic (archaeological) breeding sites identified in Childerhouse & Gales 1998 and Collins et al. 2014 that are within a 10km Euclidean (straight line) distance from a site.	To use as reference if future actions lead to proactive measures such as reintroductions.
DOC_code (DOC conservation area codes)	List of DOC conservation area codes within a site.	To query if sites are found within an area under other management priorities.
DOC_name (DOC conservation area names)	List of DOC conservation area names within a site.	See *DOC_code*.
DOC_size (DOC conservation areas size (km^2))	Total size of DOC conservation areas within a site.	To assess how much of a site is already being managed for other purposes.
DOC_pc (DOC conservation area coverage (%))	Proportion (in percent) of a site that is a DOC conservation area.	See *DOC_size*.

Contact:

Veronica F. Frans, MSc
Center for Systems Integration and Sustainability
Department of Fisheries and Wildlife
Michigan State University, USA
Email: VeroFrans@gmail.com