**Supplementary Material**

Mapping the potential distribution of the Critically Endangered Himalayan Quail *Ophrysia superciliosa* using proxy species and species distribution modelling

JONATHON C. DUNN, GRAEME M. BUCHANAN, RICHARD J. CUTHBERT, MARK J. WHITTINGHAM AND PHILIP J. K. MCGOWAN

**Contents**

Table S1. The geo-referencing accuracy of the Himalayan quail records in our database.

Table S2. The number of post-1980 Cheer Pheasant records used in our Maxent models.

Table S3. The number of post-1980 Himalayan Monal records used in our Maxent models.

Table S4. The number of Himalayan Quail records before and after vetting.

Table S5. Optimal Linear Estimation extinction dates based on vetted data.

Figure S1. Frequency polygon displaying search effort as measured by number of Galliformes records (all other species) in the locations (20km buffer) of Mussoorie and Nainital.

Table S6. WWF Ecoregions used in Maxent analysis.

Table S7. Details of covariates used in full environmental niche models for proxy species.

Table S8. The effect of different regularisation parameters (denoted as ‘Beta’) on our climate/topography models.

Table S9. The effect of different regularisation parameters (denoted as ‘Beta) on our full environmental niche models for proxy species.

Table S10. Further details for our niche models.

Table S11. The relative importance of covariates used in climate/topography models for a-b) proxy species and c) Himalayan Quail.

Table S12. The relative importance of different covariates used in our full niche models for a) Cheer Pheasant b) Himalayan Monal.

Further details of historical records used in modelling procedures

Table S1. The geo-referencing accuracy of the Himalayan quail records in our database.

|  | Pre-vetting | | Post-vetting | |
| --- | --- | --- | --- | --- |
| Geo-referencing accuracy | All | Post-1980 | All | Post-1980 |
| Accurate (1 km) | 13 | 1 | 12 | 1 |
| Close (up to 10 arc mins /18.5km) | 24 | 6 | 16 | 2 |
| Vague (10 arc mins – 48km) | 8 | 6 | 4 | 4 |
| Unknown | 10 | 4 | 2 | 1 |

Table S2. The number of post-1980 Cheer Pheasant records used in our Maxent models. The records are divided by geo-referencing accuracy.

|  |  |  |
| --- | --- | --- |
| Locational error | Number of records | % of total records |
| Accurate (1 km) | 255 | 55 |
| Close (18.5 km) | 89 | 19 |
| Vague (48 km) | 109 | 23 |
| Unknown | 16 | 3 |
| Total | 469 | 100 |

Table S3. The number of post-1980 Himalayan Monal records used in our Maxent models. The records are divided by geo-referencing accuracy.

|  |  |  |
| --- | --- | --- |
| Locational error | Number of records | % of total records |
| Accurate (1 km) | 163 | 51 |
| Close (18.5 km) | 66 | 21 |
| Vague (48 km) | 78 | 25 |
| Unknown | 10 | 3 |
| Total | 317 | 100 |

Table S4. The number of Himalayan Quail records before and after vetting. Nineteen records were omitted as they lacked latitude and longitude coordinates and a date. Two other records were also omitted that were collected post-1980 as they were 20 km from Mussoorie and Nainital and lacked information on record and observation type rendering them unreliable.

|  | Number of records | |
| --- | --- | --- |
| Data | Pre-vetting | Post-vetting |
| All | 55 | 34 |
| Sight | 29 | 17 |
| Specimen | 8 | 4 |
| Unknown | 15 | 13 |
| Heard and Seen | 1 | 0 |
| Second hand | 2 | 0 |

Further details relating to Optimal Linear Estimation technique and results

Table S5. Optimal Linear Estimation extinction dates based on vetted data. It was impossible to generate an extinction date based on the most reliable specimen data only, so we present the results when both all records and records from the last few years were used in the calculations.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Data used | Number records | Number year classes | Extinction date | Upper CI | Lower CI |
| All | 34 | 13 | 2023 | 2120 | 1999 |
| Last five years | 11 | 5 | 2010 | 2194 | 1996 |
| Specimens only | 4 | 2 | N/A | N/A | N/A |

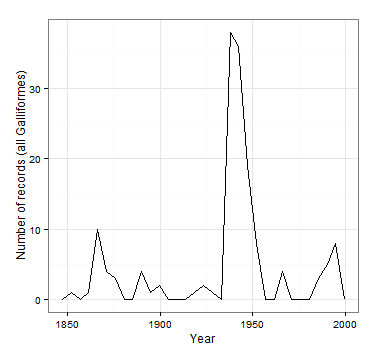


Figure S1. Frequency polygon displaying search effort as measured by number of Galliformes records (all other species) in the locations (20km buffer) of Mussoorie and Nainital. The troughs of this graph illustrate that search effort periodically drops to zero through time, thus violating one of the assumptions of the Optimal Linear Estimation method.

Further details relating to niche modelling procedures

Table S6. WWF Ecoregions used in Maxent analysis. KEY: Himalayan Quail = HQ, Himalayan Monal = HM, Cheer Pheasant = CQ

| Name | Species present |
| --- | --- |
| Baluchistan xeric woodlands | HM |
| Brahmaputra Valley semi-evergreen forests | HM |
| Eastern Himalayan alpine shrub and meadows | HM, CP |
| Eastern Himalayan broadleaf forests | HM |
| Eastern Himalayan subalpine conifer forests | HM, CP |
| Himalayan subtropical broadleaf forests | HM, CP |
| Himalayan subtropical pine forests | HM, HQ, CP |
| Northeastern Himalayan subalpine conifer forests | HM |
| Northwestern Himalayan alpine shrub and meadows | HM, CP |
| Northwestern thorn scrub forests | HM, CP |
| Nujiang Langcang Gorge alpine conifer and mixed forests | HM |
| Rock and Ice | HM, CP |
| Terai-Duar savanna and grasslands | CP |
| Upper Gangetic Plains moist deciduous forests | HM, HQ, CP |
| Western Himalayan alpine shrub and meadows | HM, CP |
| Western Himalayan broadleaf forests | HM, CP |
| Western Himalayan subalpine conifer forests | HM, CP |

Table S7. Details of covariates used in full environmental niche models for proxy species. Key: 1 = used, 0 = not used, ER = ecoregion, SS = study site, lab = labelled, part = partitioned by elevation. Note: ER and SS indicate different geographic extents used in the modelling procedure and part indicates the different ways of delimiting the Himalayan monal’s summer distribution.

|  | Species | | | | |
| --- | --- | --- | --- | --- | --- |
| Covariate | cheer ER | cheer SS | hmonal all | hmonal summer (lab) | hmonal summer (part) |
| Jan NDVI | 1 | 1 | 0 | 0 | 0 |
| Feb NDVI | 0 | 0 | 1 | 1 | 0 |
| Mar NDVI | 0 | 0 | 0 | 1 | 0 |
| Apr NDVI | 0 | 1 | 0 | 0 | 0 |
| May NDVI | 0 | 0 | 0 | 0 | 0 |
| Jun NDVI | 1 | 1 | 1 | 0 | 1 |
| Jul NDVI | 0 | 0 | 1 | 0 | 0 |
| Aug NDVI | 0 | 0 | 1 | 1 | 1 |
| Sep NDVI | 0 | 0 | 0 | 0 | 0 |
| Oct NDVI | 0 | 0 | 0 | 0 | 0 |
| Nov NDVI | 0 | 0 | 0 | 0 | 0 |
| Dec NDVI | 1 | 1 | 0 | 0 | 0 |
| Mean annual temperature | 0 | 0 | 1 | 0 | 1 |
| Mean annual variability temperature | 1 | 1 | 1 | 1 | 1 |
| Mean annual precipitation | 0 | 0 | 0 | 0 | 0 |
| Mean variation annual precipitation | 0 | 1 | 1 | 1 | 0 |
| elevation | 1 | 1 | 1 | 0 | 1 |
| Slope | 0 | 0 | 1 | 0 | 1 |
| Aspect | 0 | 0 | 0 | 0 | 0 |
| study site | 0 | 1 | 1 | 1 | 1 |
| ecoregions | 1 | 0 | 0 | 0 | 0 |
| occupied neighbour | 0 | 0 | 0 | 0 | 0 |
| Total number covariates | 6 | 8 | 10 | 6 | 7 |

Table S8. The effect of different regularisation parameters (denoted as ‘Beta’) on our climate/topography models.

| Species | Beta | Log Likelihood | Parameters | Sample Size | AIC score | AICc score | BIC score |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Cheer | 0.5 | -5557.04 | 28 | 464 | 11170.08 | 11173.81 | 11286.00 |
| 1 | -5569.21 | 31 | 464 | 11200.43 | 11205.02 | 11328.76 |
| 2 | -5592.40 | 31 | 464 | 11246.80 | 11251.40 | 11375.14 |
| 5 | -5680.40 | 22 | 464 | 11404.80 | 11407.10 | 11495.88 |
| Hmonal all | 0.5 | -3908.49 | 21 | 303 | 7858.99 | 7862.27 | 7936.97 |
| 1 | -3908.30 | 20 | 303 | 7856.60 | 7859.58 | 7930.87 |
| 2 | -3907.62 | 20 | 303 | 7855.23 | 7858.21 | 7929.51 |
| 5 | -3920.70 | 17 | 303 | 7875.41 | 7877.55 | 7938.54 |
| Hmonal summer labels | 0.5 | -121.61 | 10 | 11 | x | x | x |
| 1 | -126.18 | 6 | 11 | 264.36 | 285.36 | 266.74 |
| 2 | -128.09 | 5 | 11 | 266.17 | 278.17 | 268.16 |
| 5 | -130.43 | 2 | 11 | 264.85 | 266.35 | 265.65 |
| Hmonal summer partitioned | 0.5 | -800.17 | 9 | 66 | 1618.35 | 1621.56 | 1638.06 |
| 1 | -805.12 | 9 | 66 | 1628.24 | 1631.45 | 1647.94 |
| 2 | -806.90 | 7 | 66 | 1627.81 | 1629.74 | 1643.13 |
| 5 | -818.59 | 6 | 66 | 1649.19 | 1650.61 | 1662.33 |
| Hquail | 0.5 | -60.97 | 7 | 8 | x | x | x |
| 1 | -64.75 | 6 | 8 | 141.49 | 225.49 | 141.97 |
| 2 | -71.43 | 6 | 8 | 154.86 | 238.86 | 155.34 |
| 5 | -87.33 | 2 | 8 | 178.67 | 181.07 | 178.82 |

Table S9. The effect of different regularisation parameters (denoted as ‘Beta) on our full environmental niche models for proxy species.

| Species | Beta | Log Likelihood | Parameters | Sample Size | AIC score | AICc score | BIC score |
| --- | --- | --- | --- | --- | --- | --- | --- |
| cheer ER | 0.5 | -5584.328678 | 20 | 464 | 11208.66 | 11210.55 | 11291.46 |
| 1 | -5599.473152 | 21 | 464 | 11240.95 | 11243.04 | 11327.88 |
| 2 | -5623.476049 | 21 | 464 | 11288.95 | 11291.04 | 11375.89 |
| 5 | -5701.703703 | 16 | 464 | 11435.41 | 11436.62 | 11501.65 |
| cheer SS | 0.5 | -5598.818513 | 24 | 464 | 11245.64 | 11248.37 | 11344.99 |
| 1 | -5598.818513 | 24 | 464 | 11245.64 | 11248.37 | 11344.99 |
| 2 | -5603.208521 | 21 | 464 | 11248.42 | 11250.51 | 11335.35 |
| 5 | -5657.378093 | 15 | 464 | 11344.76 | 11345.83 | 11406.85 |
| hmonal all | 0.5 | -3860.417004 | 32 | 303 | 7784.834 | 7792.656 | 7903.673 |
| 1 | -3860.417004 | 32 | 303 | 7784.834 | 7792.656 | 7903.673 |
| 2 | -3863.302875 | 31 | 303 | 7788.606 | 7795.927 | 7903.731 |
| 5 | -3886.056443 | 14 | 303 | 7800.113 | 7801.571 | 7852.105 |
| hmonal summer (lab) | 0.5 | -120.4607464 | 7 | 11 | 254.9215 | 292.2548 | 257.7068 |
| 1 | -121.8299976 | 5 | 11 | 253.66 | 265.66 | 255.6495 |
| 2 | -123.400783 | 5 | 11 | 256.8016 | 268.8016 | 258.791 |
| 5 | -129.1854223 | 4 | 11 | 266.3708 | 273.0375 | 267.9624 |
| hmonal summer (part) | 0.5 | -796.4174687 | 10 | 66 | 1612.835 | 1616.835 | 1634.731 |
| 1 | -799.860979 | 8 | 66 | 1615.722 | 1618.248 | 1633.239 |
| 2 | -804.532957 | 7 | 66 | 1623.066 | 1624.997 | 1638.393 |
| 5 | -822.9198906 | 5 | 66 | 1655.84 | 1656.84 | 1666.788 |

Table S10. Further details for our niche models. Note the number of unique records is lower than the number of records available. This is because Maxent automatically omits spatial duplicates of records. Feature function = the shape of the response curves for each model explanatory variable. Key: ER = ecoregions, SS = study site, lab = labelled, part = partitioned by elevation, ON = occupied neighbour, l = linear, q = quadratic, p = product.

| Species | Model | Number unique records used | Study site delimitation method | Feature function | Regularisation value | Mean threshold value |
| --- | --- | --- | --- | --- | --- | --- |
| Hquail | Climate | 5 | ON | l | 5 | 0.715 |
| Cheer ER | Full | 192 | ER | lqp | 0.5 | 0.345 |
| Cheer SS | Full | 192 | SS | lqp | 1 | 0.305 |
| Hmonal all | Full | 216 | SS | lqp | 1 | 0.365 |
| Hmonal summer (lab) | Full | 10 | SS | lp | 1 | 0.548 |
| Hmonal summer (part) | Full | 39 | SS | lp | 0.5 | 0.308 |

Table S11. The relative importance of covariates used in climate/topography models for a-b) proxy species and c) Himalayan Quail. Percent contribution is calculated as follows: in each iteration of the training algorithm, the increase in regularised gain is added to the contribution of the corresponding variable. Permutation importance is calculated as follows: for each variable in turn, the values of that variable of training presence and background data are randomly permutated. The model is then re-evaluated on the permuted data and the resulting drop in AUC is shown, normalised to a percentage.

|  |  |  |
| --- | --- | --- |
| 1. Cheer Pheasant climate model | | |
| Variable | Percent contribution | Permutation importance |
| ecoregions | 70.2 | 36 |
| bio\_4 | 17 | 17.5 |
| bio\_15 | 3.8 | 4.1 |
| bio\_1 | 2.1 | 14.1 |
| elevation | 2 | 23 |
| bio\_12 | 1.7 | 3.2 |
| aspect | 1.6 | 0.5 |
| slope | 1.5 | 1.7 |

|  |  |  |
| --- | --- | --- |
| B. Himalayan Monal climate model | | |
| Variable | Percent contribution | Permutation importance |
| bio\_15 | 43.1 | 5.8 |
| elevation | 22.1 | 33.6 |
| bio\_1 | 13 | 23.1 |
| bio\_12 | 7.7 | 15 |
| bio\_4 | 6 | 15.4 |
| slope | 5.1 | 4.5 |
| aspect | 3 | 2.7 |
| ssite | 0 | 0 |

|  |  |  |
| --- | --- | --- |
| C. Himalayan Quail climate model | | |
| Variable | Percent contribution | Permutation importance |
| bio\_4 | 66.8 | 91.9 |
| bio\_1 | 23.4 | 1.6 |
| elevation | 9 | 4.4 |
| bio\_12 | 0.8 | 2.1 |
| slope | 0 | 0 |
| himqu\_m | 0 | 0 |
| bio\_15 | 0 | 0 |
| aspect | 0 | 0 |

Table S12. The relative importance of different covariates used in our full niche models for a) Cheer Pheasant b) Himalayan Monal. Percent contribution is calculated as follows: in each iteration of the training algorithm, the increase in regularised gain is added to the contribution of the corresponding variable. Permutation importance is calculated as follows: for each variable in turn, the values of that variable of training presence and background data are randomly permutated. The model is then re-evaluated on the permuted data and the resulting drop in AUC is shown, normalised to a percentage.

|  |  |  |
| --- | --- | --- |
| A. Cheer Pheasant full model | | |
| Variable | Percent contribution | Permutation importance |
| ecoregions | 59.2 | 51.2 |
| bio\_4 | 15.9 | 25.3 |
| jun\_mean | 11.1 | 2.6 |
| jan\_mean | 5.5 | 8.9 |
| dec\_mean | 5.2 | 6.9 |
| elevation | 3.1 | 5.1 |

|  |  |  |
| --- | --- | --- |
| B. Himalayan Monal full model | | |
| Variable | Percent contribution | Permutation importance |
| bio\_15 | 22.7 | 3.4 |
| jun\_mean | 14.1 | 0.4 |
| jul\_mean | 13.8 | 3.9 |
| elevation | 10.1 | 9.3 |
| feb\_mean | 9.5 | 33 |
| bio\_1 | 8 | 25.8 |
| bio\_4 | 7.6 | 1.3 |
| slope | 7.2 | 3.8 |
| aug\_mean | 7.1 | 19.2 |
| ssite\_hquail | 0 | 0 |