4 Discussion

This project comprised four distinct elements: the estimation of ‘flower score’ and its value for modelling twite distribution; the description of the key data sets and examination of any obvious relationships between different land cover and land use categories with twite distribution using the zonal statistics as table analysis; the development of a range of predictive models based on environmental and land use data; and testing of the best models on independent data. This section considers these elements and discusses how the results might be useful in predicting the distribution of twite across Britain, what restrictions and limitations might apply to interpreting the results and ways in which the work might be improved or taken further.

4.1 ‘Flower Score’ Estimation

In terms of modelling the twite’s feeding habitat preferences, and hence the species’ distribution, the survey of ‘Flower Score’ was of limited value. There are no obvious relationships between the abundance of the food plants as recorded and twite numbers. It is very likely that this is an artefact of the survey method, and perhaps the way in which the birds forage. This survey involved recording floristic diversity of certain key species at specific locations where twite had been recorded previously. However, it is probable that twite will forage over considerable areas and move between sites depending on the presence or absence of these food plants, and indeed some less important species too. Hence, an overall higher flower score may not actually make a site more attractive to twite if, for example, it is isolated from other sites. Thus the dynamics of twite foraging will be very complex and further thought is required in designing an appropriate measure of suitability for twite foraging, perhaps on a wider scale and less site specific.

While these results were of limited value to this exercise there remains the possibility that further work in this area might yield an insight into how twite interact with their forage areas and why birds nest at locations that are sometimes a considerable distance from the best forage areas. Perhaps nest sites are traditional and the gaps between colonies represent the reduction in numbers over a long period of time. It may be more likely that there is a critical issue of optimising nest locations so that birds can forage over a number of sites at different and perhaps well dispersed locations at different times throughout the breeding season.

4.2 Zonal Statistics as Table Analysis

The zonal statistics analysis proved to be very useful in shaping the subsequent development of the models. This analysis identified the most common land cover types and average percentage cover of each in tetrads where twite were recorded and compared that with the average percentage cover of each of these categories nationally. Three categories made up the vast majority of land cover in twite tetrads (Figure 21) and each displayed an obvious difference in twite tetrads compared to the national average. Interestingly, there were some differences in the results for 1990 and 2000, but the same three categories, *Moor & heathlands*, *Natural grasslands*, and *Pastures*, were the dominant cover types. The first two might be expected to be important habitat types for twite, but *Pasture* is not necessarily ideal for them. That being said, it is important to note that while *Pasture* was a major land cover type in twite tetrads the coverage was noticeably lower than the national average, suggesting that *Pasture* is not a favourable habitat for twite.

In Scotland the LCS88 data showed a different picture, although *Heather moor* was the dominant cover type *Bogs & peatlands* was second, far higher than the national average. Also, *Arable* coverage was very low in twite tetrads compared with the national average and *Coarse grasslands* was also less common. Otherwise, most of the categories varied a little but not hugely from the national average. These results indicate potential habitats that might play a key role in predicting twite presence, either because they are common in twite areas or the contrary. Thus categories like *Pasture* and *Arable* are more likely to be useful in predicting where twite will not be found. The results of this analysis proved to be very useful in terms of selecting the variables that ought to be used in developing the model, i.e. those that displayed a distinct pattern for twite tetrads compared with the national picture.

In practical terms of model development, this analytical method was very valuable in that it allowed the amalgamation of data at varying resolutions to be combined at a common scale and on a common reference system, based on spatial location. GIS played a key role in this analysis as it would not have been readily possible to bring the land cover and Agricultural Census data into a common format otherwise. Similarly, the elevation and climate data was imported at different resolutions and this was merged into the same format using the spatial join function that GIS provides, where the mean value of cells making up a tetrad was attached to each tetrad.

4.3 Model Development

There are many varied methods used to develop predictive models, some more complex than others. In many respects this was a relatively simple approach to model building, being largely based upon logistic regression analysis involving a wide selection of environmental and land use parameters. However, the models developed here include some additional elements that include spatial relationships between different parameters as well. The use of logistic regression as the foundation of all the models allows easy assessment of the models' relative performance using standard statistical measures of 'Goodness of Fit'.

All of the models appear to provide excellent predictions in terms of the general pattern of twite distribution, as illustrated in the probability maps in Appendix 6 and as measured by the area under the ROC curve graphs in Appendix 7. The lowest ROC value achieved by any of the models for probability of twite being present was 0.788, with the majority being in excess of 0.9. These results suggest that the models should predict twite distribution excellently. However, when the ROC values for the predicted distribution is derived the results are less impressive, producing obviously poor distribution maps and much lower ROC values. The lowest ROC value was 0.548 for the *Basic Model plus Elevation* which is little better than the results that might be expected by a random guess. Only one model (*LCS88*) produced a ROC value in excess of 0.7 which is typically used as a benchmark of acceptability.

So why are the results for the probability of finding twite so good by comparison with the predicted distributions? The fact that the models are heavily biased towards predicting a negative result, i.e. because there are vastly more tetrads without twite than with them the chance of finding twite in a tetrad chosen at random is very low, only 3.2 % for Britain. Hence, the measured performance of the models is skewed towards higher values of ROC since in the majority of cases the models predict a negative answer, and simply by the odds this is more likely to be correct than not. This does not sound very positive in terms of the quality of these models, but it is important to consider how many tetrads with twite are predicted correctly. In essence a random guess would have a 3.2 % chance of actually finding a tetrad with twite present. The models however all have a far better strike rate with the best British one (*Geographic Zones model*) being in excess of 30 %. The best Scottish result being in excess of 40 % (Table 13) compared to a random chance of c.10 % in Scotland. These results effectively indicate that the best model for Britain provides a 10x better chance of predicting the distribution of twite than a random guess, and the Scottish model a 4x better chance.

There are various factors that might impose limitations on the effectiveness of these models and their potential future use or development. The models, being developed through a process of logistic regression, are heavily dependent upon the original twite data against which all the other parameters were assessed. While the dataset is a very impressive one, containing what is still the best overall record of the distribution of twite around Britain, it does have some limitations. Most notable is the fact that the survey did not cover the whole of the country, although every effort was made to ensure that it was as representative as possible, there were gaps. The data was also collected by a wide and varied group of individuals, who would by definition have differing abilities. The twite is not a particularly impressively coloured or obvious bird, often being recognised by its call rather than its appearance. Hence there is a genuine possibility that the species is under-recorded throughout Britain, but it is impossible to assess the extent to which this is true. However, it is obvious from a closer examination of the distribution map (Figure 6) that there have been differing rates of survey coverage across Britain as illustrated in the extract from that map (Figure 35) which shows an unusual arrangement of tetrads in Caithness where twite were recorded but they were strangely absent from most of the tetrads in the surrounding area. Thus under-recording and variable recording might be an issue that could have a knock-on effect on the strength of the models.

The Agricultural Census data is based upon annual submissions by farmers and crofters across the country. However, the level and scale of their activities will greatly affect whether or not a farmer submits a return. In the crofting areas, where many twite are found, the majority of crofters probably do not submit annual returns as they may not be claiming subsidies. Hence, there is perhaps an issue of reliability for this data set for some parts of the country. Added to that, the Corine Land Cover data is derived from satellite imagery at a European scale and has only very broad categories. Some of these categories are very relevant to twite distribution but some are patently too broad, e.g. *Moors and heathlands*. It is difficult to have a great deal of belief that there will be a strong correlation between twite presence and the extent of moorland if that moorland can include wet acid heath dominated by *Molinia* grass and rushes, and other areas might be dominated by dry heather yet be included in the same category. The LCS data does offer some improvement in this regard.

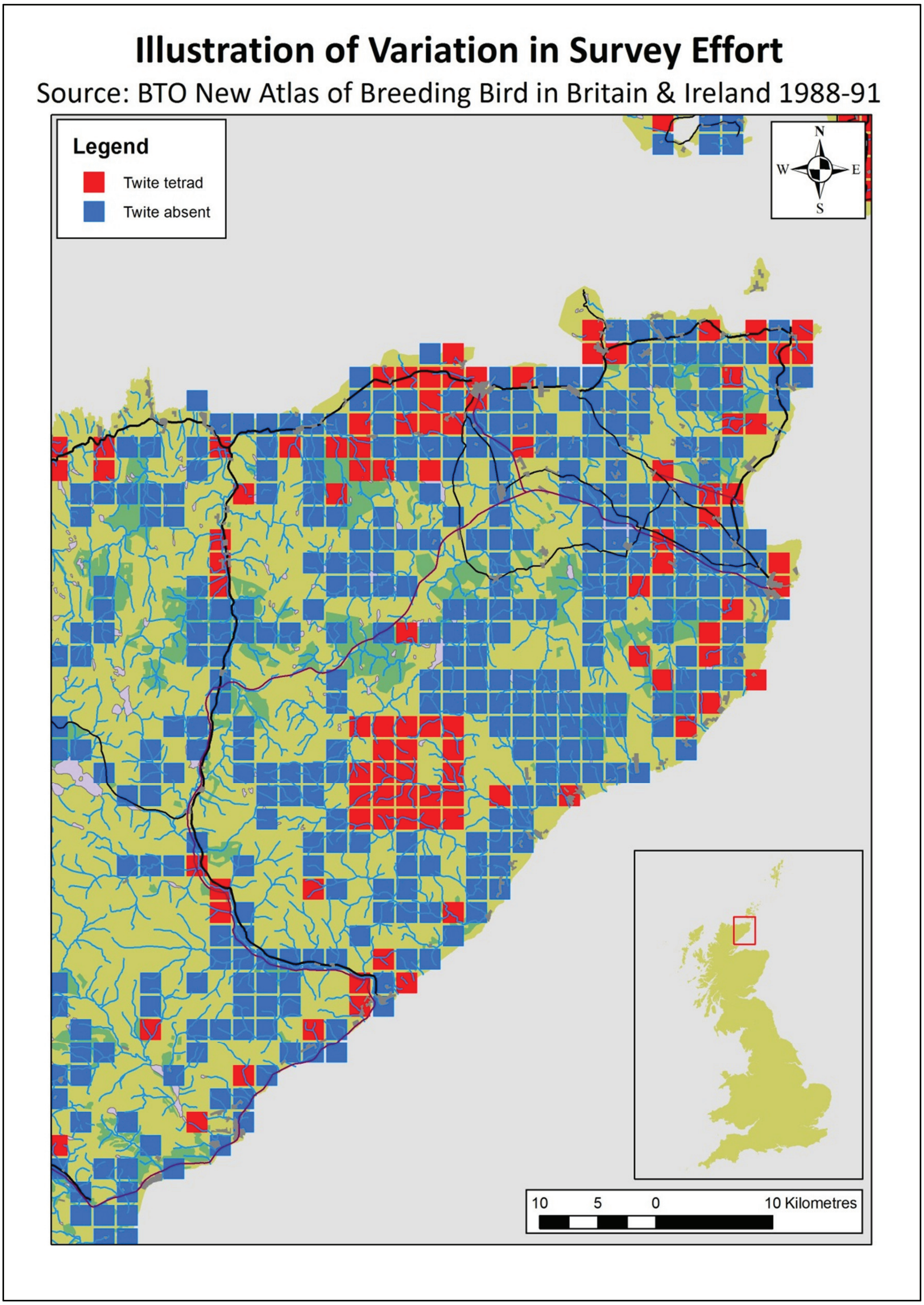

Figure 35 - Illustration of the variation is survey effort in the BTO88-91 dataset

If the results are considered visually in terms of how accurately the models predict actual presence, for example in North and South Uist (Figure 36), the Geographic Zones model appears to correctly predict the majority of twite tetrads quite well especially along the west coast at lower elevations but it is far from being accurate all the time. If the Elevation Zones model is examined more closely, for example in the Pennines (Figure 37) it is obvious that the twite are constrained largely to the higher ground and the model performs reasonably well, especially in predicting no twite in the surrounding lower lying areas. This accurate prediction of absence is common to all of the models, no doubt in part due to the fact that twite are largely absent throughout Britain, so statistically it is more likely to be true.

4.4 Testing the Models against Independent Data

The results of testing the models against data from 1999-2000 were disappointing in so far as the ROC values were very low, and in most cases little better than a random chance. The perfomrance of the models relative to each other was consistent with the 1988-91 data but none-the-less only the *Geographic Zones model* produced a ROC value approaching an acceptable level. This can perhaps be explained partly by the relatively small sample size, only 1 % of the size of the model development data set. Also, the condition that the test data should be consistent with the original data may be failed on account of the Agricultural Census data formats had changed slightly and not all of the variables could be guaranteed to be exactly comparable across the country. However, the fact that the same relative performance was achieved highlights the best models more clearly.

Interestingly, in all cases the reduction in ROC values for distribution compared with probability of twite was much less for the model testing data. Perhaps this is due to the higher proportion of twite squares (c. 30 % as opposed to 3.2 %). This may also explain why the model performs relatively poorly, given it was derived to predict twite presence on the assumption that there would be very few twite tetrads. In retrospect a alternative approach might have been to use a sub-set of the original data for testing, excluded from the model development, and with similar proportions of tetrads with twite.

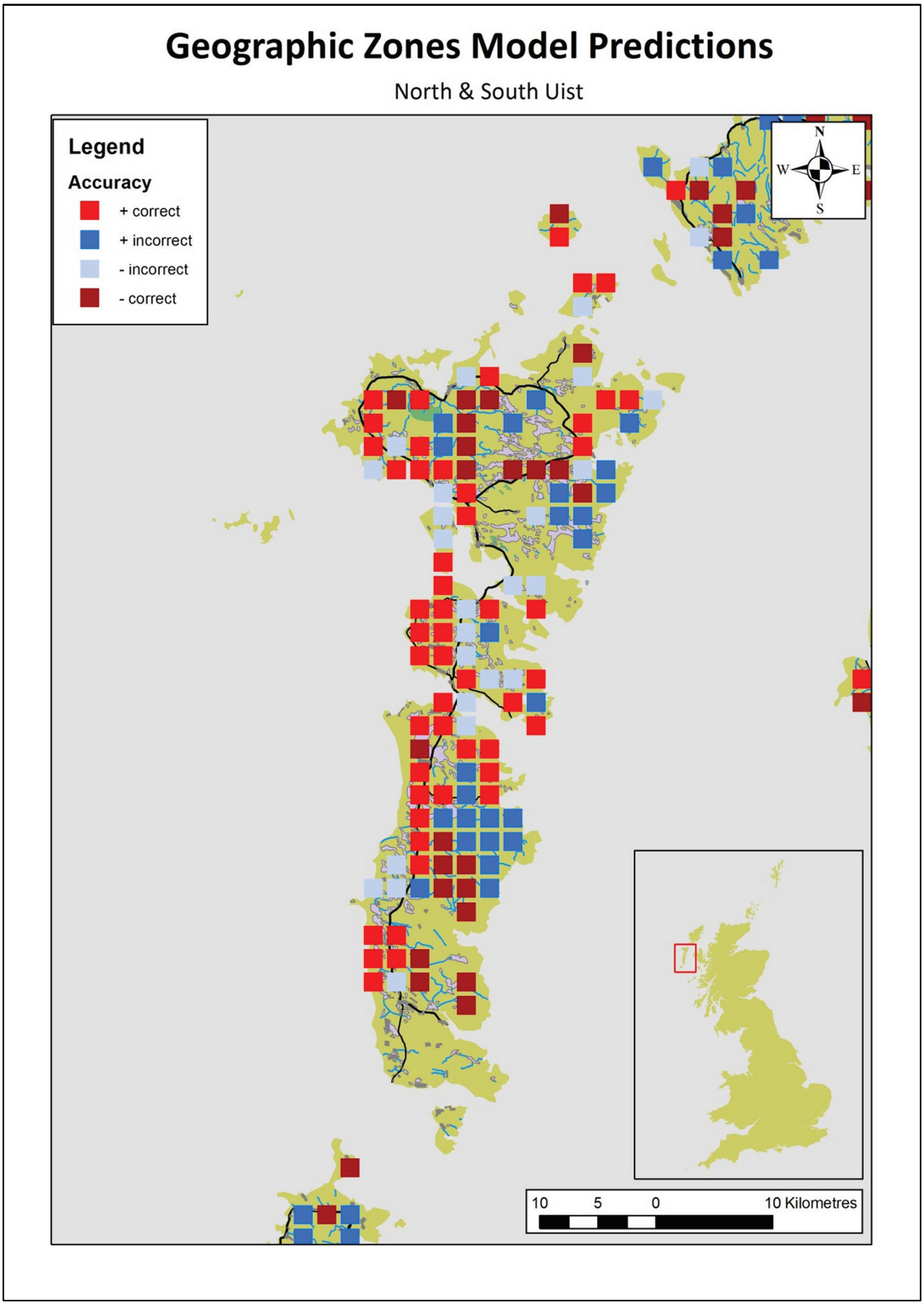

Figure 36 - Geographic Zones Model accuracy for North & South Uist

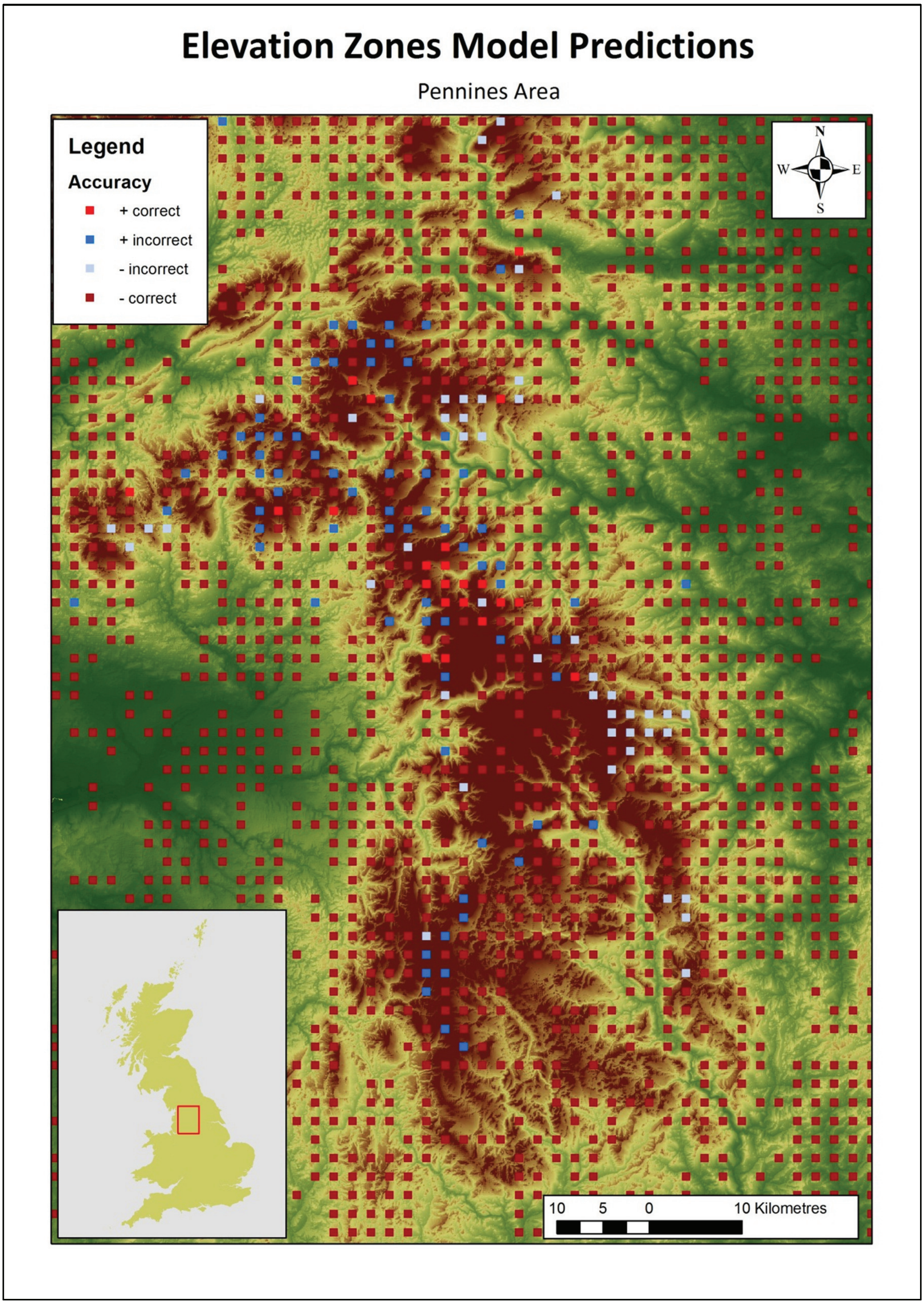

Figure 37 - Elevation Zones Model accuracy for the Pennines

This questionable performance of the models begs the question, how could this modelling approach be improved to produce a more reliable local scale prediction of twite distribution?

4.5 Potential Improvements to the Models

Attempts were made to introduce elements of spatial interaction between the different parameters in some of the models. Initially by applying geographic zones, then elevation zones, the idea that relationships might be inconsistent across the country was introduced. Later on the concept of distance between nest sites and forage areas was built into the models using the *Habitat Proximity* weighting factor, giving greater chance of twite being predicted in areas where *Moorland & heathland* was within 1.6 km of *Natural grasslands* or *Pasture*. Perhaps by extending these spatially oriented approaches to modelling to include Geographically Weighted Regression (GWR) it might be possible to improve the estimation of these interactions and variations (Fotheringham et al 2002, Jetz et al 2005, Osborne et al 2007). GWR was considered for this project but was ruled out on the basis of time constraints and the lack of a GWR tool that undertakes logistic regression within ArcGIS. It is possible to do linear or multiple regression using the ArcGIS GWR tool but this would not be effective for a dichotomous outcome such as presence or absence. That being said, a GWR approach to modelling twite distribution might prove very useful, if an appropriate scale of kernel could be applied, and data was available at that resolution. This would have the advantage of removing any concerns about arbitrary boundaries for geographic zones or elevation zones as well.

The wider scale distribution of twite appears superficially to be modelled reasonably well by the models developed here, but more locally it is very patchy. The best models appear to be predicting the absence of twite very well but on closer inspection of areas where twite ought to be found the predictions are less reliable. This might be due to a need for more detailed local analysis of the factors affecting distribution. Twite may utilise several neighbouring tetrads in a season for nesting and foraging, purely by dint of the fact that tetrad boundaries are regular and arbitrary in terms of twite behaviour. Thus it might be concluded that a small scale model such as derived here might be useful for broad impressions of distribution but it is unlikely to be able to predict accurately the more localised distribution of this species.

Added to the question of scale is the issue of data quality. The data used here was defined primarily by what was readily available for the correct time periods and without excessive cost or delay. The land cover data and Agricultural census data was not ideal in that few of the variables were directly applicable to modelling twite distributions. More detailed assessment of habitats and land cover might greatly improve the performance of the models, especially if heather dominated moorland and enclosed farmland with associated 'flower scores' could be obtained. Also, inclusion of a better measure of exposure, rather than simply temperature and rainfall, might be useful and could lead to improvement as this is probably a more relevant factor. The use of habitat surveys as opposed to land use and land cover data might be worth considering. This would introduce a greater degree of detail perhaps, depending on the type of survey, e.g. Phase 1 Habitat Survey or National Vegetation Classification survey. If this kind of data was amalgamated with some estimate of an areas general attractiveness to twite, via some kind of 'flower score' the results might be much improved. However, this would require a significant investment to obtain the data and would likely be cost prohibitive.