

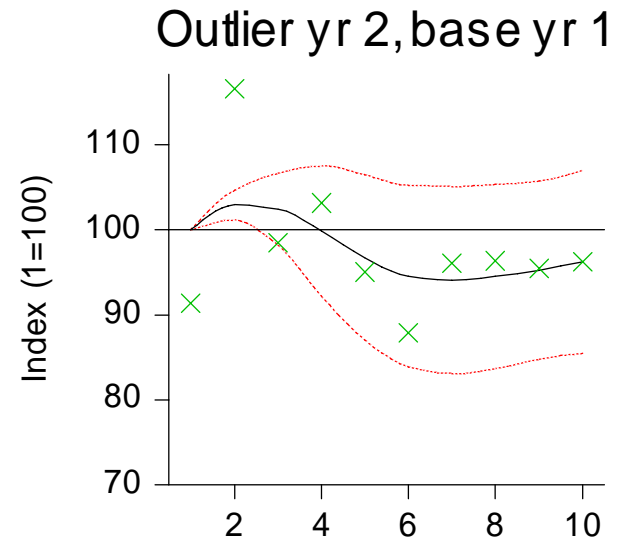
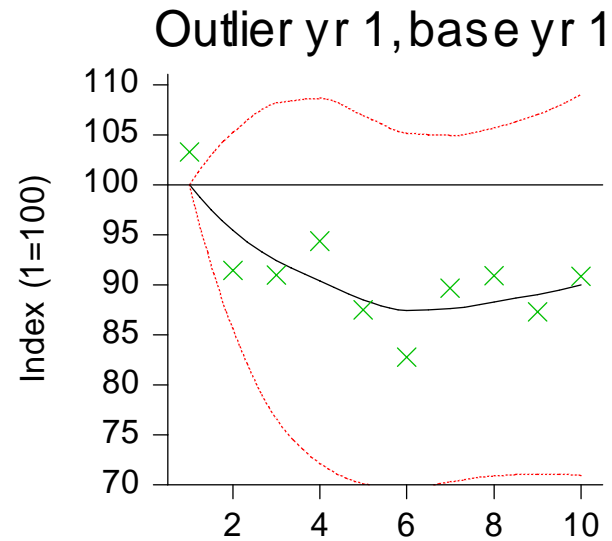
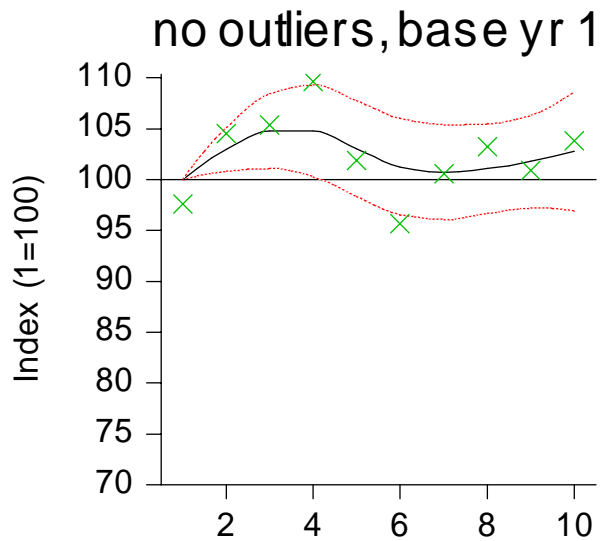
# Outliers in GAM models

- GAM method fits smoothed curve to temporal trend
- Little influence by outliers in intermediate years
- Outliers in first or last year have greater impact

# Example and simulations

- Simulated Poisson data with no trend
- 10 year, 50 sites
- Pattern of missing values added from real data
- Outliers added by taking the highest count in appropriate year and doubling
- Using either year 1 or year 2 as base year

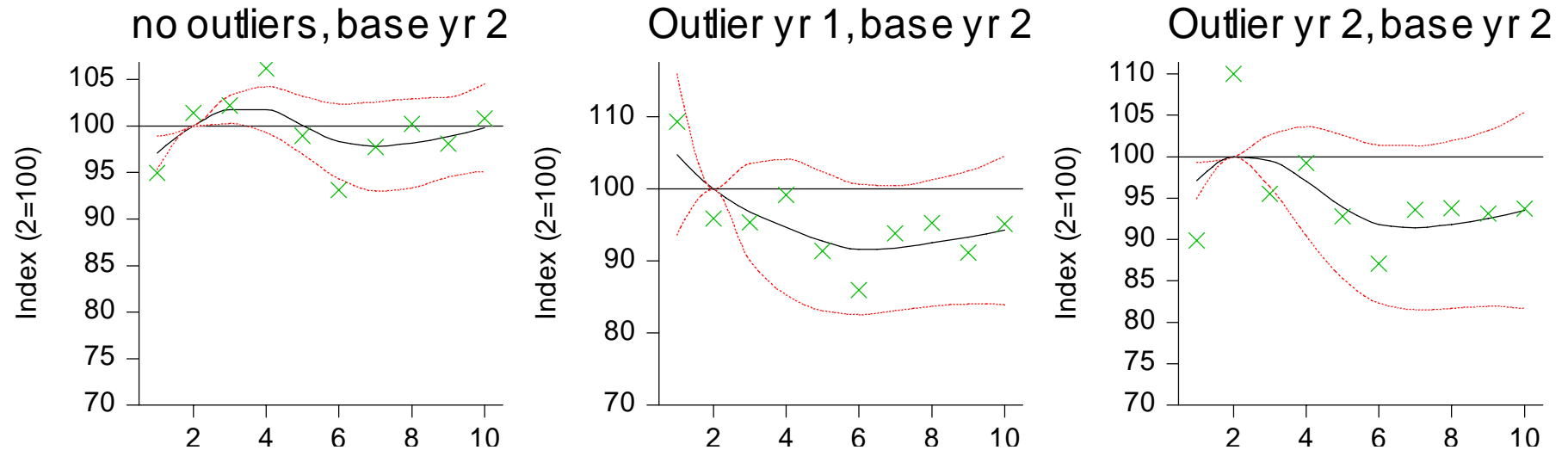
# Example – impact of outlier using year 1 as base



Points show annual estimates fitting year as a factor, black line is fitted GAM

Index at year 10 is around 103 in this example without outliers. An outlier in year 1 makes it drop to 90, or an outlier in year 2 to 96.

# Example – impact of outlier using year 2 as base



Points show annual estimates fitting year as a factor, black line is fitted GAM.

Without outliers index at year 10 is almost exactly 100. An outlier in year 1 drops it to 94 and an outlier in year 2 to 93.5.

2,000 sets of simulated data confirm the pattern shown in this example, i.e. using the second year as the base reduces the impact of outliers.

# Outliers - conclusions

- Where outliers are present in year 1, best to use year 2 as base year to avoid them having too great an impact.
- Further simulation indicates that using year 2 as base gives improved precision, even if outliers are not present.
- Similar conclusions probably apply to final year of survey – final year's data should either not be published, or should be treated as provisional.
- See *index methods for bat populations.doc* for more information.