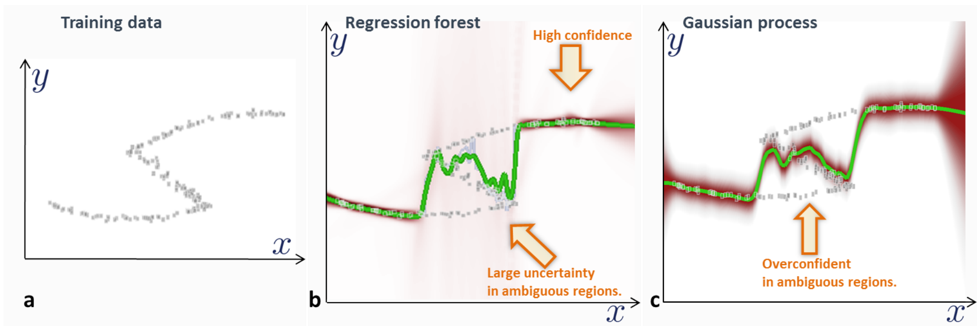


Fig. 4.10: **Comparing forests and GP on ambiguous training data.** **(a)** Input labelled training points. The data is ambiguous because a given input $x$ may correspond to multiple values of $y$. **(b)** The posterior $p(y|x)$ computed via random regression forest. The middle (ambiguous) region remains associated with high uncertainty (in grey). **(c)** The posterior computed via Gaussian Processes. Conventional GP models do not seem flexible enough to capture spatially varying noise in training points. This yields an over-confident prediction in the central region. In all these experiments the GP parameters have been automatically optimized for optimal results, using the provided Matlab code.

figs. 4.9b,c. There, we can observe how the forest can capture bi-modal distributions in the gaps (see orange arrows). Due to their piece-wise nature the regression forest seems more apt at capturing multi-modal behaviour in testing regions and thus modeling intrinsic ambiguity (different $y$ values may be associated with the same $x$ input). In contrast, the posterior of a Gaussian process is by construction a (uni-modal) Gaussian, which may be a limitation in some applications. The same uni-modal limitation also applies to the recent "relevance voxel machine" technique in [76].

This difference between the two models in the presence of ambiguities is tested further in fig. 4.10. Here the training data itself is arranged in an ambiguous way, as a "non-function" relation (see also [63] for computer vision examples). For the same value of $x$ there may be multiple training points with different values of $y$.

The corresponding testing posteriors are shown for the two models