Some residual analysis methods for space-time point processes.

Rick Paik Schoenberg, UCLA Statistics
(thanks to: Robert Clements, Andrew Bray, Ka Wong)
http://www.stat.ucla.edu/~frederic/papers1.html

1) Pixel based residuals.

2) Superthinned residuals.

3) Voronoi residuals.

Not discussed here:
numerical summaries (L-test, N-test, etc.),
functional summaries (error diagrams, weighted K-functions, etc.)
Sometimes graphical residual methods are preferable.
Suppose we have a model for $\lambda(t,x,y)$, the conditional rate at $(t,x,y)$, given previous seismicity.

1. **Pixel-based residuals.**
   Compare $N(A_i)$ with $\int_A \lambda(t, x) \, dt \, dx$, on pixels $A_i$. (Baddeley, Turner, Møller, Hazelton, 2005)

   **Problems:**

   * If pixels are large, lose power.
   * If pixels are small, residuals are mostly $\sim 0,1$.
     -- non-normality after standardization.
   * Smoothing reveals only gross features.
-- Given two competing models, can consider the difference between residuals over each pixel.

Problem: Hard to interpret. If difference = 3, is this because model A overestimated by 3? Or because model B underestimated by 3? Or because model A overestimated by 1 and model B underestimated by 2?

-- Better: consider difference between log-likelihoods, in each pixel. The result may be called deviance residuals, in analogy with residuals from logistic regression and other generalized linear models. Basically the same as information gain.

Deviance residuals effectively highlight differences between models, even with sparse data.
Fig. 4. Left panel (a): deviance residuals for model A versus C. Sum of deviance residuals is 86.427. Right panel (b): deviance residuals for model B versus C. Sum of deviance residuals is −7.468.

Fig. 3. Left panel (a): deviance residuals for model A versus B. Sum of deviance residuals is 84.393. Right panel (b): close-up of deviance residuals for model A versus B near the Imperial fault.
2. Superthinning.

**Thinning:** Suppose \( \inf \lambda(t_i, x_i, y_i) = b \).

Keep each point \((t_i, x_i, y_i)\) with probability \( \frac{b}{\lambda(t_i, x_i, y_i)} \).
Superposition: Suppose \( \sup \lambda(t, x, y) = c \).

Superpose \( N \) with a simulated Poisson process of rate \( c - \lambda(t, x, y) \).

Problems with thinning and superposition:

Thinning: Low power. If \( b = \inf \lambda(t_i, x_i, y_i) \) is small, will end up with very few points.

Superposition: Low power if \( c = \sup \lambda(t_i, x_i, y_i) \) is large: most of the residual points will be simulated.

Superthinning: superpose where \( \lambda(t, x, y) < c \), and thin where \( \lambda(t_i, x_i, y_i) > c \).
Check the resulting plot for homogeneity.
Voronoi Tessellation: (Okabe, 2000) a partitioning of $S$ into $n$ convex polygons (tiles)

$$D_i = \{ x \in X : \|x - x_i\| \leq \|x - x_j\|, \forall j \neq i \}$$

Voronoi Residual:

$$\hat{r}_i = \frac{N(D_i) - \int_{D_i} \hat{\lambda}(x_i) \, dx_i}{SE \left( \int_{D_i} \hat{\lambda}(x_i) \, dx_i \right)} = \frac{1 - |D_i| \bar{\lambda}}{SE \left( |D_i| \bar{\lambda} \right)}$$
Properties of Voronoi Residuals

- Non-parametric
- Spatially adaptive
- $|D|\bar{\lambda} \sim Gamma(3.5, 3.5)$ for homogeneous (Tanemura 2003)
- Approximately Gamma for inhomogeneous (Barr and Schoenberg 2010, Barr and Diez in progress)
- generating model = fitted model = $100x^2|y|$
- less-skewed residuals with narrower range
overprediction

underprediction
Conclusions.

* Deviance residuals are useful for comparing models on grid cells.

* Superthinned residuals do not rely on a grid and can be useful to highlight where a model overpredicts or underpredicts.

* Voronoi residuals use an automatically adaptive grid and seem great for both comparison and to see where a particular model overpredicts or underpredicts.
3. Error Diagrams

Plot (normalized) number of alarms vs. (normalized) number of false negatives (failures to predict). (Molchan 1990; Molchan 1997; Zaliapin & Molchan 2004; Kagan 2009).

Similar to ROC curves (Swets 1973).

Problems:
-- Must focus near axes.
[consider relative to given model (Kagan 2009)]
-- Difficult to see where model fits poorly.
4. Residuals: rescaling, thinning, superposing


Suppose $N$ is simple. Rescale one coordinate: move each point

$$\{t_i, x_i\} \text{ to } \{t_i, \int_0^{x_i} \lambda(t_i, x) \, dx\} \quad [\text{or to } \{\int_0^{t_i} \lambda(t, x_i) \, dt, x_i\}] .$$

Then the resulting process is stationary Poisson.

Problems:
* Irregular boundary, plotting.
* Points in transformed space hard to interpret.
* For highly clustered processes: boundary effects, loss of power.