

# **At Variance: The Cluster Sampling Paradigm**

*Geoffrey Embree, Northeast Utilities*

## **Summary**

Many energy efficiency evaluations examine multiple measure installations per site, but treat that site as a single observation for purposes of estimating variance. Rather than treating each site as an individual observation, cluster sampling treats each independently evaluated measure as a separate data point, reducing the variance of the final estimate. In the past, strong intra-cluster correlation has limited the effectiveness of a clustered approach, but as CFLs and other technologies have become more mainstream, cluster sampling may more accurately represent the margin of error in a typical saturation study or home energy audit, and explains the impressive stability of these estimates in recent years.

This poster presents some simple calculations for a model lighting saturation study supporting an actual variance equal to approximately half of the per-site variance estimate, implying an effective sample size equal to four times the number of sites. Evaluations are already realizing the benefit of cluster sampling approaches – the only thing left is to quantify the improvement and convey to end-users the enhanced precision delivered by comprehensive site visits.