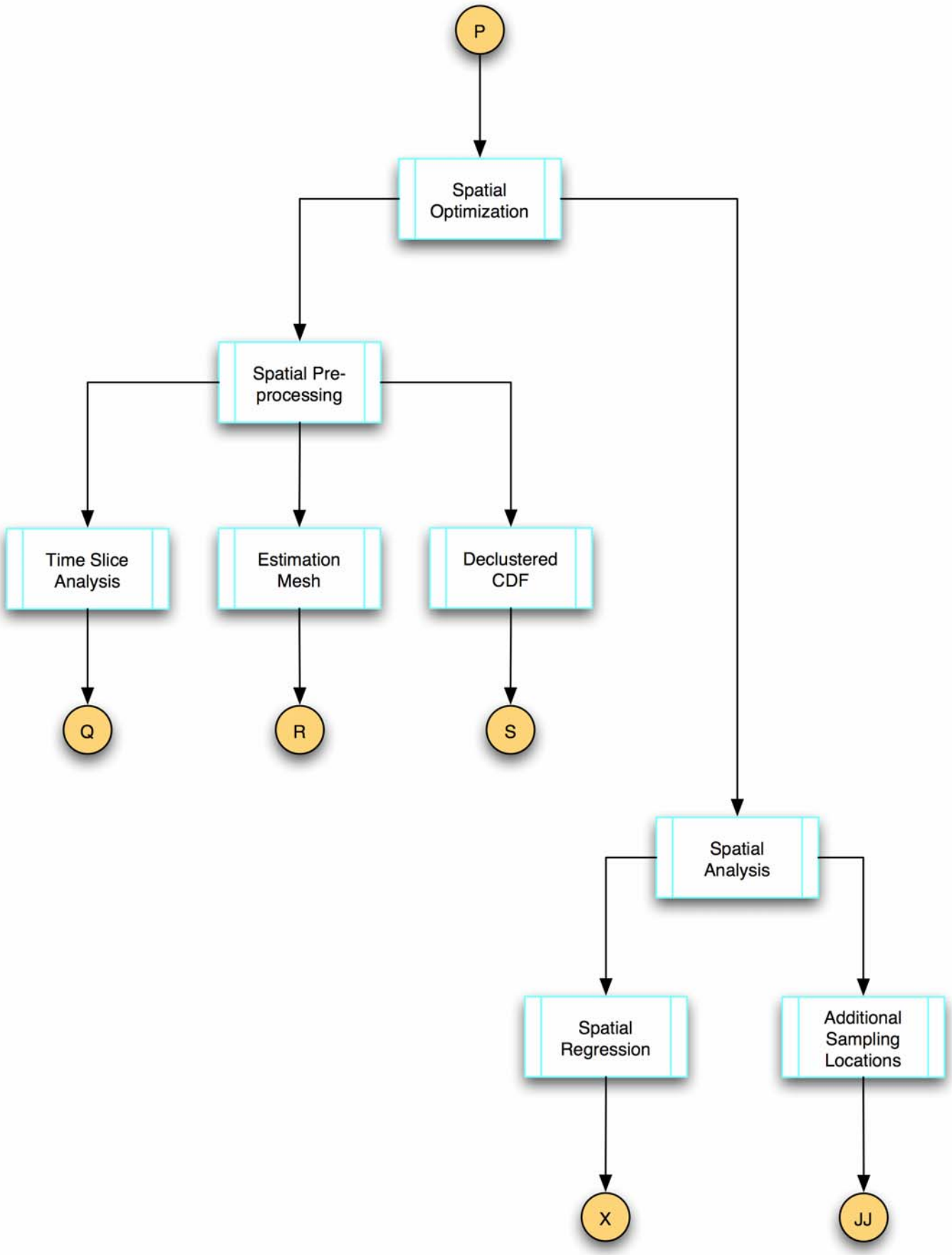
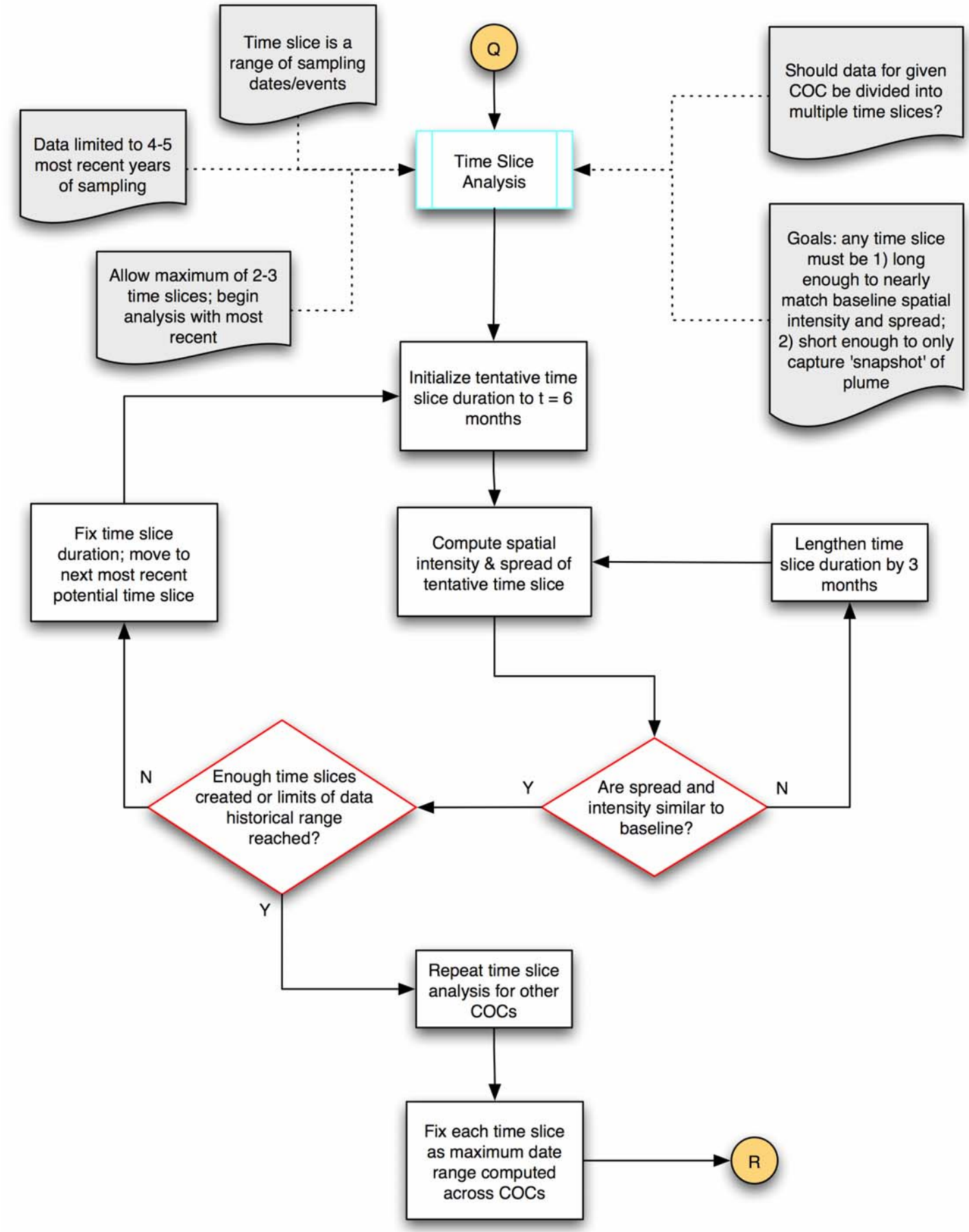
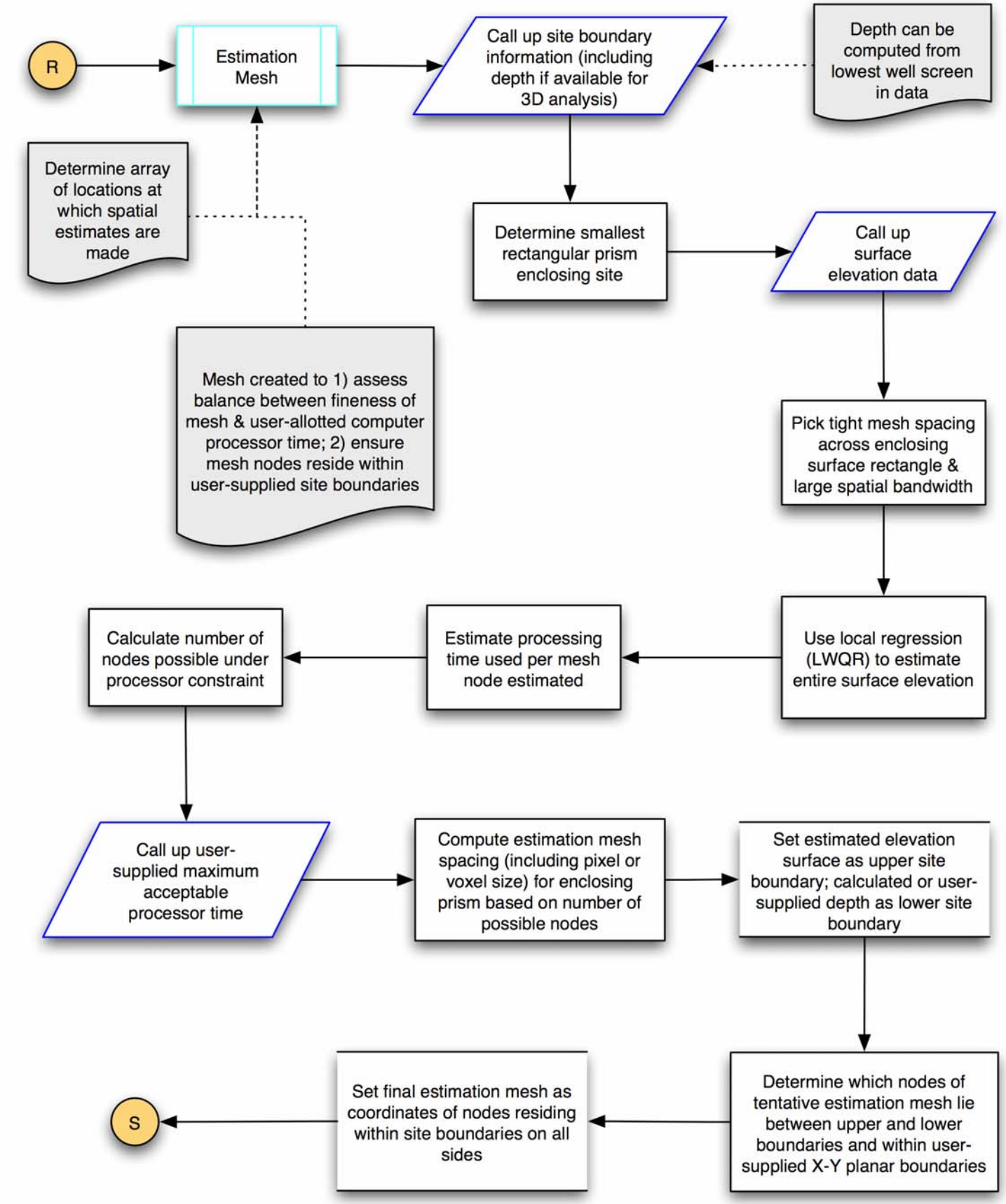
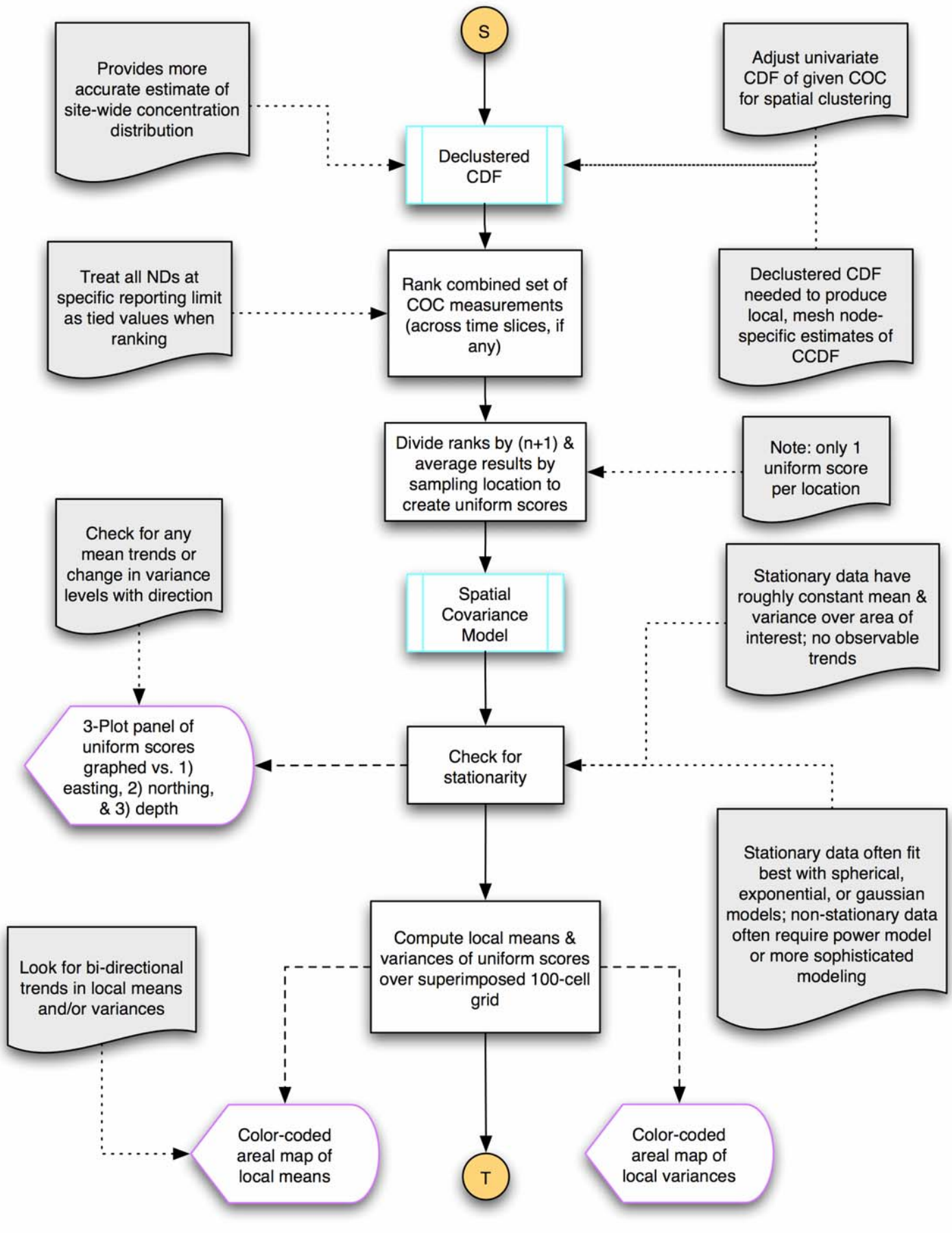
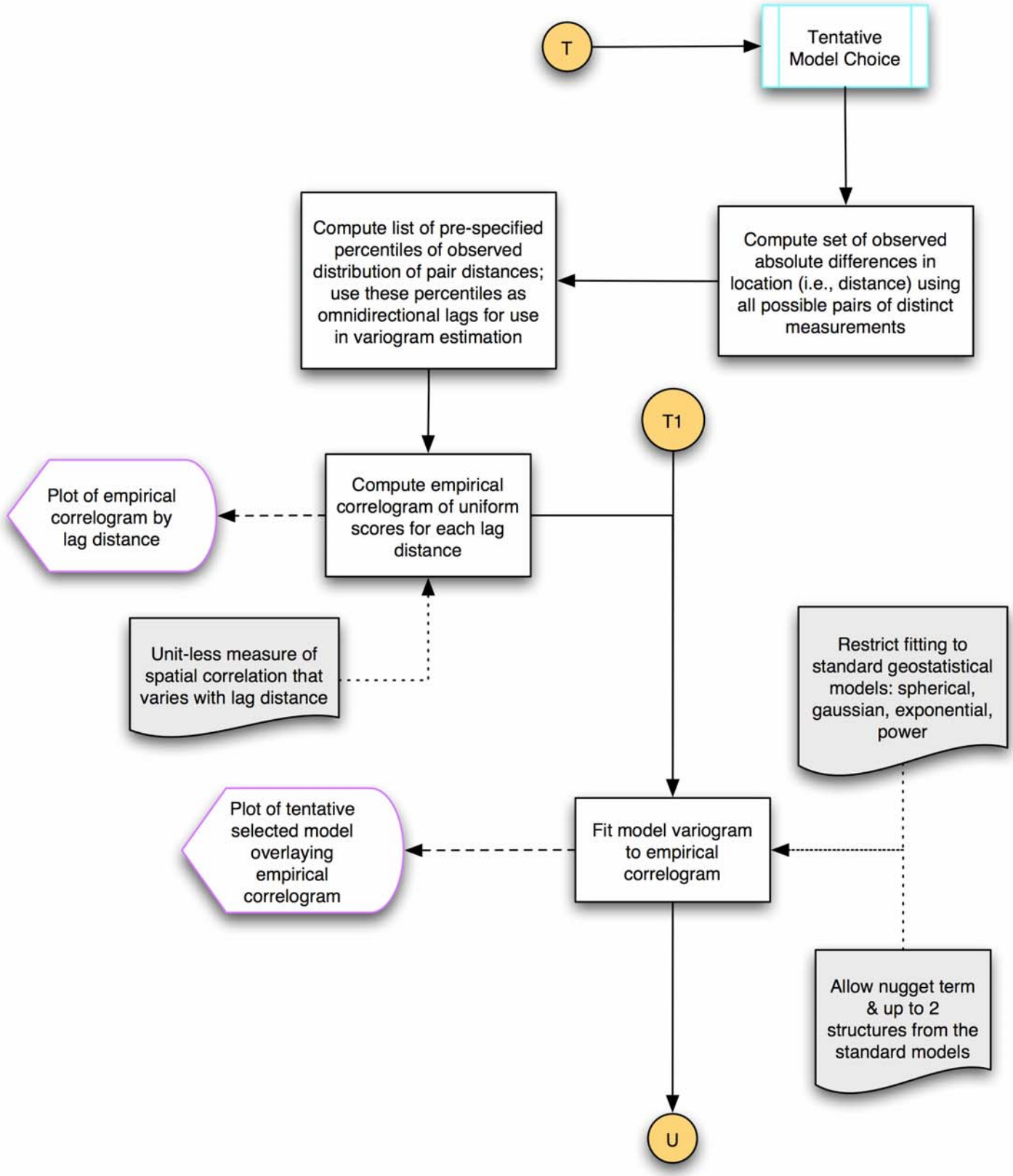
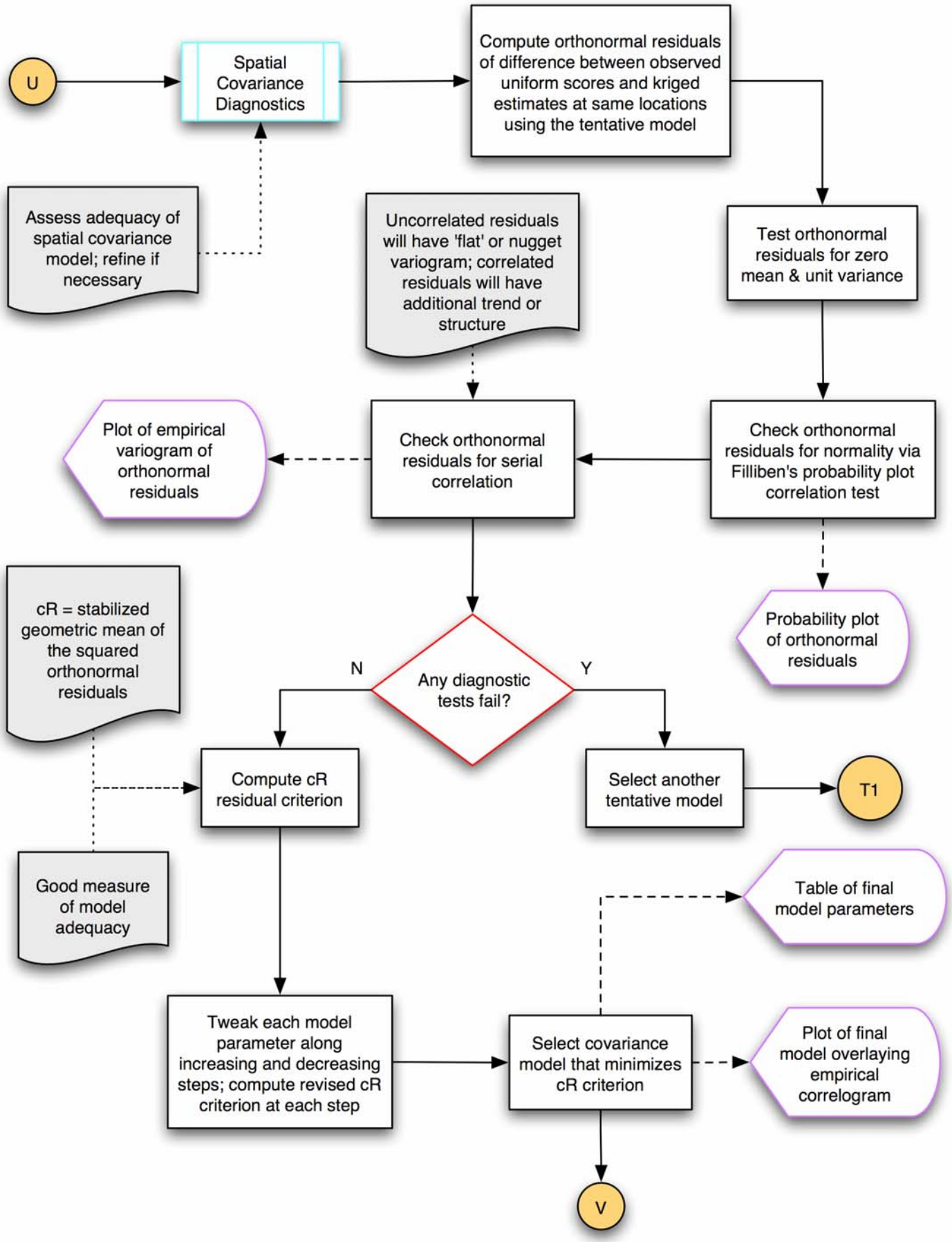

P
Spatial Optimization
Spatial Pre-processing
Time Slice Analysis
Estimation Mesh
Declustered CDF
Q
R
S
Spatial Analysis
Spatial Regression
Additional Sampling Locations
X
JJ

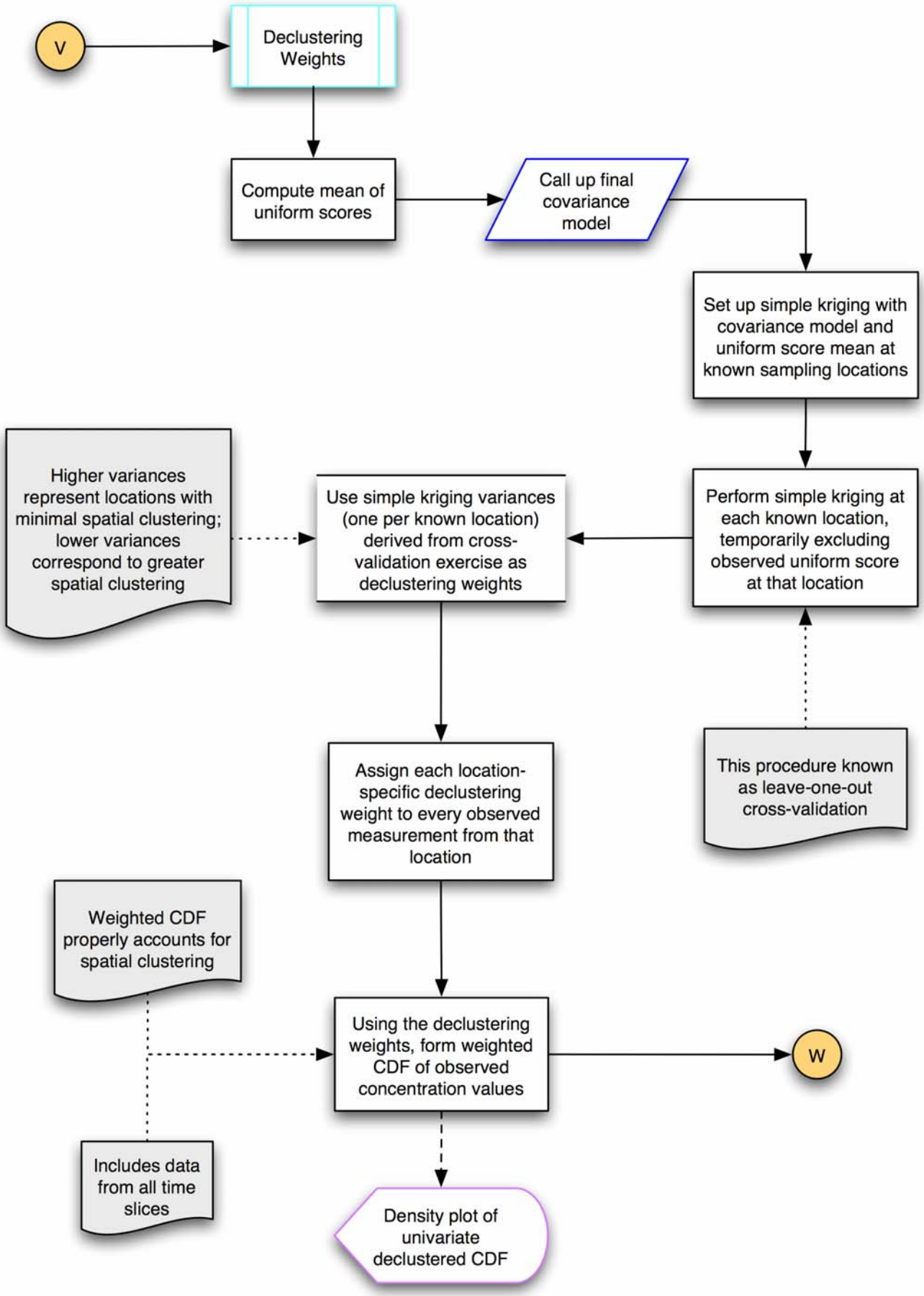

Q
Time slice is a range of sampling dates/events
Should data for given COC be divided into multiple time slices?
Data limited to 4-5 most recent years of sampling
Time Slice Analysis
Allow maximum of 2-3 time slices; begin analysis with most recent
Goals: any time slice must be 1) long enough to nearly match baseline spatial intensity and spread; 2) short enough to only capture 'snapshot' of plume
Initialize tentative time slice duration to t = 6 months
Fix time slice duration; move to next most recent potential time slice
Compute spatial intensity & spread of tentative time slice
Lengthen time slice duration by 3 months
N
Enough time slices created or limits of data historical range reached?
Y
Are spread and intensity similar to baseline?
N
Y
Repeat time slice analysis for other COCs
Fix each time slice as maximum date range computed across COCs
R

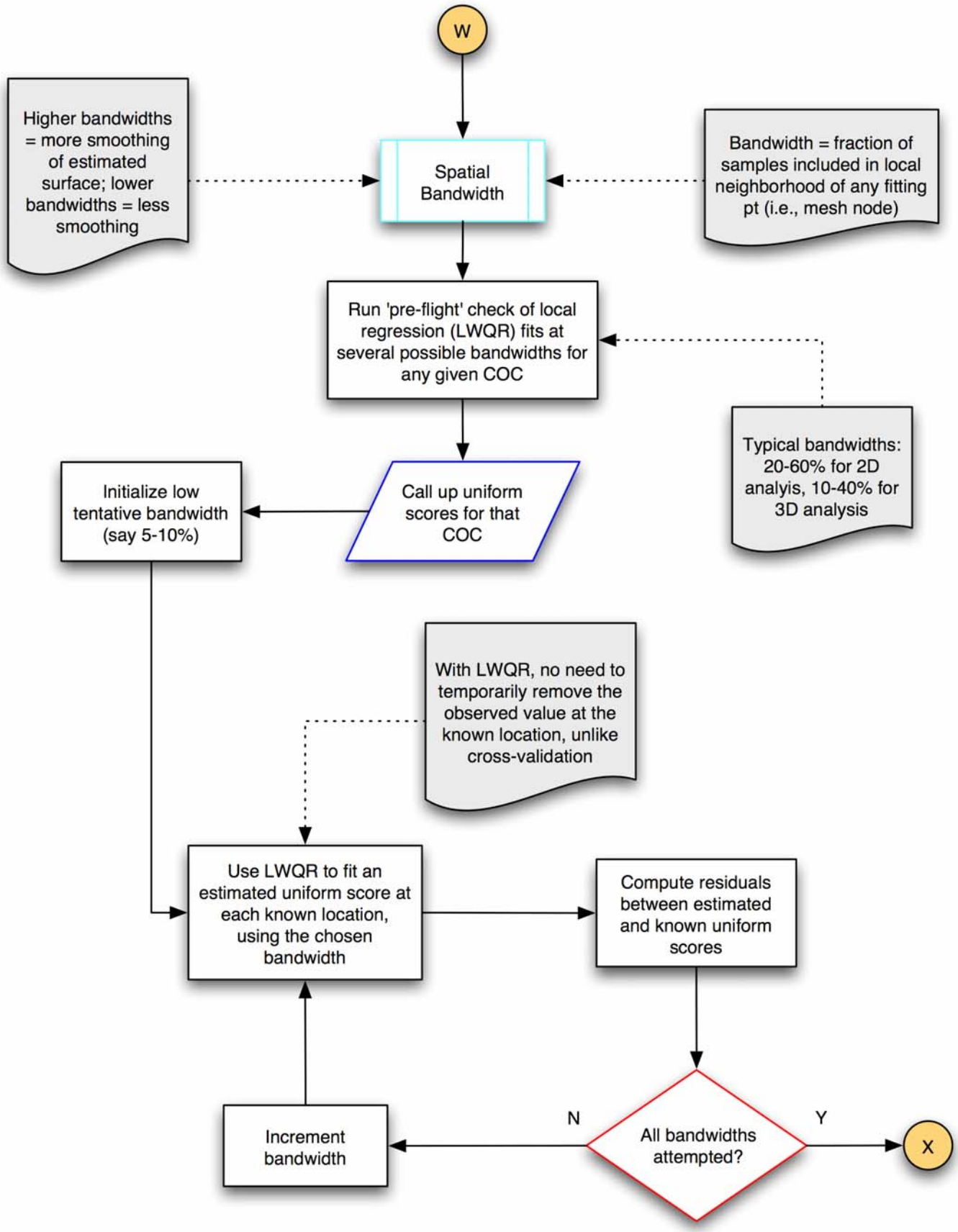

R
Estimation Mesh
Call up site boundary information (including depth if available for 3D analysis)
Depth can be computed from lowest well screen in data
Determine array of locations at which spatial estimates are made
Determine smallest rectangular prism enclosing site
Call up surface elevation data
Mesh created to 1) assess balance between fineness of mesh & user-allotted computer processor time; 2) ensure mesh nodes reside within user-supplied site boundaries
Pick tight mesh spacing across enclosing surface rectangle & large spatial bandwidth
Calculate number of nodes possible under processor constraint
Estimate processing time used per mesh node estimated
Use local regression (LWQR) to estimate entire surface elevation
Call up user-supplied maximum acceptable processor time
Compute estimation mesh spacing (including pixel or voxel size) for enclosing prism based on number of possible nodes
Set estimated elevation surface as upper site boundary; calculated or user-supplied depth as lower site boundary
S
Set final estimation mesh as coordinates of nodes residing within site boundaries on all sides
Determine which nodes of tentative estimation mesh lie between upper and lower boundaries and within user-supplied X-Y planar boundaries


S
Provides more accurate estimate of site-wide concentration distribution
Adjust univariate CDF of given COC for spatial clustering
Declustered CDF
Treat all NDs at specific reporting limit as tied values when ranking
Rank combined set of COC measurements (across time slices, if any)
Declustered CDF needed to produce local, mesh node-specific estimates of CCDF
Divide ranks by (n+1) & average results by sampling location to create uniform scores
Note: only 1 uniform score per location
Check for any mean trends or change in variance levels with direction
Spatial Covariance Model
Stationary data have roughly constant mean & variance over area of interest; no observable trends
3-Plot panel of uniform scores graphed vs. 1) easting, 2) northing, & 3) depth
Check for stationarity
Stationary data often fit best with spherical, exponential, or gaussian models; non-stationary data often require power model or more sophisticated modeling
Compute local means & variances of uniform scores over superimposed 100-cell grid
Look for bi-directional trends in local means and/or variances
Color-coded areal map of local means
T
Color-coded areal map of local variances


T
Tentative Model Choice
Compute set of observed absolute differences in location (i.e., distance) using all possible pairs of distinct measurements
Compute list of pre-specified percentiles of observed distribution of pair distances; use these percentiles as omnidirectional lags for use in variogram estimation
T1
Compute empirical correlogram of uniform scores for each lag distance
Plot of empirical correlogram by lag distance
Unit-less measure of spatial correlation that varies with lag distance
Restrict fitting to standard geostatistical models: spherical, gaussian, exponential, power
Fit model variogram to empirical correlogram
Plot of tentative selected model overlaying empirical correlogram
Allow nugget term & up to 2 structures from the standard models
U


U
Spatial Covariance Diagnostics
Compute orthonormal residuals of difference between observed uniform scores and kriged estimates at same locations using the tentative model
Assess adequacy of spatial covariance model; refine if necessary
Uncorrelated residuals will have 'flat' or nugget variogram; correlated residuals will have additional trend or structure
Test orthonormal residuals for zero mean & unit variance
Plot of empirical variogram of orthonormal residuals
Check orthonormal residuals for serial correlation
Check orthonormal residuals for normality via Filliben's probability plot correlation test
Probability plot of orthonormal residuals
cR = stabilized geometric mean of the squared orthonormal residuals
N
Any diagnostic tests fail?
Y
Compute cR residual criterion
Select another tentative model
T1
Good measure of model adequacy
Table of final model parameters
Tweak each model parameter along increasing and decreasing steps; compute revised cR criterion at each step
Select covariance model that minimizes cR criterion
Plot of final model overlaying empirical correlogram
V


V
Declustering Weights
Compute mean of uniform scores
Call up final covariance model
Set up simple kriging with covariance model and uniform score mean at known sampling locations
Perform simple kriging at each known location, temporarily excluding observed uniform score at that location
Use simple kriging variances (one per known location) derived from cross-validation exercise as declustering weights
Higher variances represent locations with minimal spatial clustering; lower variances correspond to greater spatial clustering
This procedure known as leave-one-out cross-validation
Assign each location-specific declustering weight to every observed measurement from that location
Weighted CDF properly accounts for spatial clustering
Using the declustering weights, form weighted CDF of observed concentration values
W
Includes data from all time slices
Density plot of univariate declustered CDF


W
Higher bandwidths = more smoothing of estimated surface; lower bandwidths = less smoothing
Spatial Bandwidth
Bandwidth = fraction of samples included in local neighborhood of any fitting pt (i.e., mesh node)
Run 'pre-flight' check of local regression (LWQR) fits at several possible bandwidths for any given COC
Typical bandwidths: 20-60% for 2D analyis, 10-40% for 3D analysis
Initialize low tentative bandwidth (say 5-10%)
Call up uniform scores for that COC
With LWQR, no need to temporarily remove the observed value at the known location, unlike cross-validation
Use LWQR to fit an estimated uniform score at each known location, using the chosen bandwidth
Compute residuals between estimated and known uniform scores
Increment bandwidth
N
All bandwidths attempted?
Y
X