DART_LAB Tutorial Section 4: Other Updates for an Observed Variable.
‘Classical’ Monte Carlo algorithm by Evensen. Note: earliest references have error, use caution.
First, fit a gaussian to the ensemble sample.
Obtain observation and observation error distribution.
Generate a random draw from the observation likelihood.
Associate it with the first sample of the prior ensemble.
Associate a random draw from the observation likelihood with each prior ensemble member. This is called *generating perturbed observations*.
Associate a random draw from the observation likelihood with each prior ensemble member. This is called
*generating perturbed observations.*
Associate a random draw from the observation likelihood with each prior ensemble member. This is called *generating perturbed observations*. 
Associate a random draw from the observation likelihood with each prior ensemble member. This is called generating perturbed observations.
We now have a sample of the joint distribution of the prior mean and observation.
The Ensemble Kalman Filter (Perturbed Observations)

Adjusting the mean of the observation sample helps. Adjusting the variance to be exact may also help (or not). Outliers are a potential problem but could be removed.
For each prior/observation pair, find the mean of the posterior distribution.
Prior sample standard deviation measures uncertainty of prior mean estimate.
Observation likelihood standard deviation measures the uncertainty of the observation estimate.
Take the product of the prior and observation distributions for the first sample. This is the standard product of gaussians.
Mean of the product is a random sample of the posterior. Product of random samples is random sample of product.
Repeat this operation for every pair of prior and observation.
Repeat this operation for every pair of prior and observation.
Repeat this operation for every pair of prior and observation.
Repeat this operation for every pair of prior and observation.
Posterior sample retains much of prior samples structure; this is more apparent for larger ensembles. Posterior sample mean and variance converge as a function of the ensemble size.
Apply forward operator to each ensemble member.
Get prior ensemble in observation space.
Step 1: Get continuous prior distribution density.

- Place \((\text{ens}_\text{size} + 1)^{-1}\) mass between adjacent ensemble members.
- Reminiscent of rank histogram evaluation method.
Step 1: Get continuous prior distribution density.

- Place $(\text{ens\_size} + 1)^{-1}$ mass between adjacent ensemble members.
- Reminiscent of rank histogram evaluation method.
Step 1: Get continuous prior distribution density.

- Place \((\text{ens\_size} + 1)^{-1}\) mass between adjacent ensemble members.
- Reminiscent of rank histogram evaluation method.
Step 1: Get continuous prior distribution density.
- Place \((\text{ens}_\text{size} + 1)^{-1}\) mass between adjacent ensemble members.
- Reminiscent of rank histogram evaluation method.
Step 1: Get continuous prior distribution density.

- Partial gaussian kernels on tails, $N(tail\_mean, ens\_sd)$.
- \textit{tail\_mean selected so that (ens\_size + 1)^{-1} mass is in tail.}
Step 2: Use **likelihood** to compute weight for each ensemble member.

- Analogous to classical particle filter.
- Can be extended to non-gaussian obs. likelihoods.
Step 2: Use **likelihood** to compute weight for each ensemble member.
- Can approximate interior likelihood with linear fit; for efficiency.
Step 3: Compute continuous posterior distribution.

- Approximate likelihood with trapezoidal quadrature, take product.

(Displayed product normalized to make posterior a PDF).
Step 3: Compute continuous posterior distribution.

- Approximate likelihood with trapezoidal quadrature, take product.
  (Displayed product normalized to make posterior a PDF).
Step 3: Compute continuous posterior distribution.

- Approximate likelihood with trapezoidal quadrature, take product.
  (Displayed product normalized to make posterior a PDF).
Step 3: Compute continuous posterior distribution.

- Approximate likelihood with trapezoidal quadrature, take product.

(Displayed product normalized to make posterior a PDF).
Step 3: Compute continuous posterior distribution.

- Product of prior gaussian kernel with likelihood for tails.
- Easy for gaussian likelihood.
- More quadrature if non-Gaussian likelihood.
Step 4: Compute updated ensemble members:

- $(\text{ens}_\text{size} + 1)^{-1}$ of posterior mass between each ensemble pair.
- $(\text{ens}_\text{size} + 1)^{-1}$ in each tail.
- Uninformative observation has no impact.
Compare to standard Ensemble Adjustment Filter (EAKF).

Nearly gaussian case, differences are small.
Rank Histogram gets rid of outlier that is clearly inconsistent with obs.
EAKF can’t get rid of outlier.
Large prior variance from outlier causes EAKF to shift all members too much towards observation.
Rank Histogram handles multimodal prior and compelling observation.

EAKF still bimodal; left mode is inconsistent with everything. Lorenz_63 can have priors like this.
Convective-scale models (and land models) have analogous behavior. Convection may fire at ‘random’ locations. Subset of ensembles will be in right place, rest in wrong place. Want to aggressively eliminate convection in wrong place.
Matlab exercises `oned_ensemble`, `twod_ensemble`, `oned_model`, `run_lorenz_63` and `run_lorenz_96` all allow selection of EnKF or RHF for assimilation.

In `oned_ensemble` and `twod_ensemble`, be sure to try the EnKF repeatedly. It’s a stochastic algorithm so it produces a different answer each time.

That’s all Folks!