A land data assimilation system (LDAS) merges observations (or satellite retrievals) of land surface hydrological conditions, including soil moisture, snow, and terrestrial water storage (TWS), into a numerical model of land surface processes. In theory, the output from such a system is superior to estimates based on the observations or the model alone, thereby enhancing our ability to understand, monitor, and predict key elements of the terrestrial water cycle. In practice, however, several conceptual problems can interfere with realizing the potential improvements from data assimilation. Of particular concern is the frequent mismatch between the assimilated observations and the land surface model variables of interest. The seminar will discuss recent research with the ensemble-based NASA GEOS-5 LDAS to address various aspects of this mismatch. These aspects include (i) the assimilation of coarse-scale observations into higher-resolution land surface models, (ii) the partitioning of satellite observations (such as TWS retrievals) into their constituent water cycle components, (iii) the forward modeling of microwave brightness temperatures over land for radiance-based land surface data assimilation, and (iv) the selection of the most relevant types of observations for the analysis of a specific water cycle variable (such as root zone soil moisture). At its core, the solution to the above challenges involves the careful construction of an observation operator that maps from the land surface model variables of interest to the space of the assimilated observations.