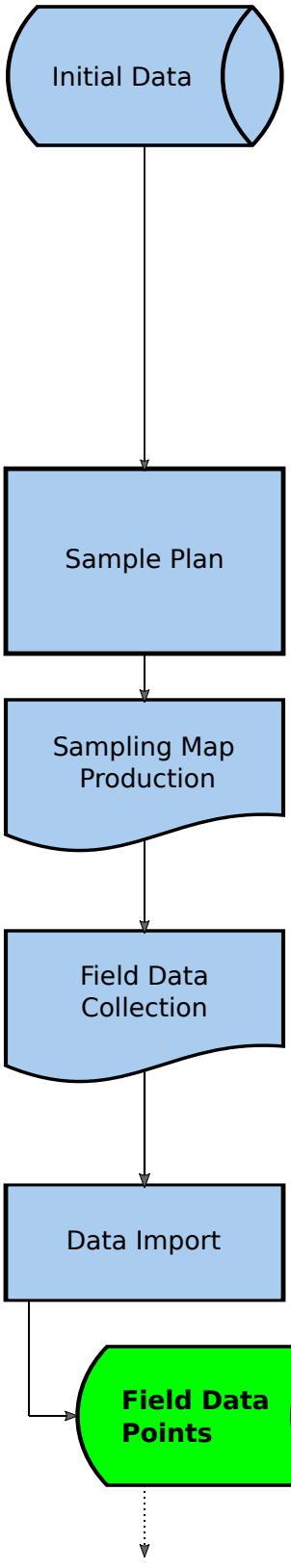
BC's Predictive Ecosystem Mapping Process

Part A: Field Sampling



Steps	Details	BEC DevExchange	pemgeneratr

Initial Data	- Digital Terrain Model (25m^2)
	
	- Defined area of interest (snapped to 100m grid)
	- BEC Subzones
	- BCData sources:
	  - Freshwater atlas (lakes, wetlands, and rivers)
	  - roads
	  - VRI: age class and conifer leading
	  - Harvested Areas of BC
	  - Forest Tenure Harvesting Authority Polygons
	  - BC Major Cities Points
	  - Fire Perimeters - Current
	  - Fire Perimeters - Historical
	  - TANTALIS - Surface Ownership
		00_folder_set_up.Rmd
		
		01_BaseLayerConsolidate.Rmd
			`aoi_snap()`
			`plan_get_baselayers()` PENDING

Sample Plan	1. Generate sample plan covariates 2. Generate sample cost surface (raster) 3. Sliced Conditioned Latin Hyper-Cube sample points and associated survey transects.	01_raster_BaseLayers.Rmd 03a_Stage1_SampleDesign.Rmd	`plan_covariates()` PENDING `plan_exclusion()` PENDING `plan_costsurface()` PENDING `plan_locatesampling()` PENDING `plan_transects()` PENDING

Sampling Map Production	- Using QGIS -- ATLAS module
	  - 1:100,000
	  - 1:2,500
	- Export to ipad Avenza's 'PDF Map' as a '*map collection*'
	- Sampling *Schema* set for project
		Manual Processing	Manual Processing

Field Data Collection	- Use high precision (sub-metre) GPS - Complete all fields as per the schema - Record track-log - Record locations where Site Series changes. - Record site series you are entering. Looking ahead, not behind - Export as GPX and Shapefile	Manual Processing	Manual Processing

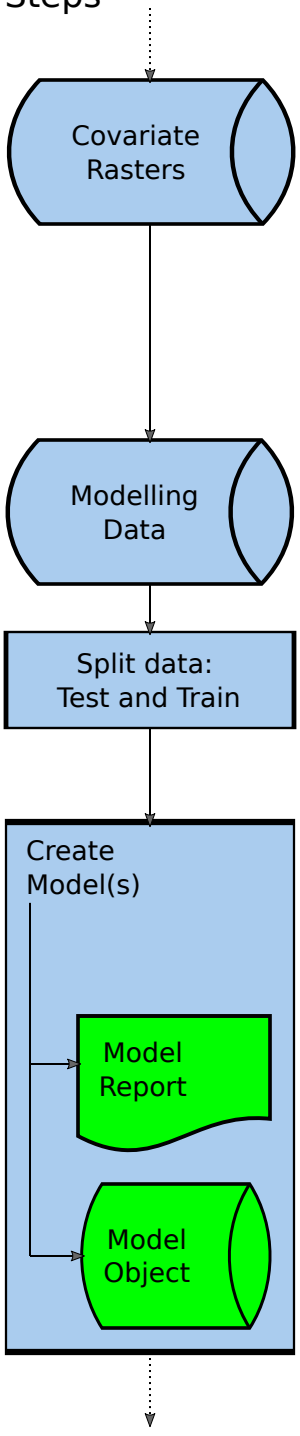
Data Import	- imports GPX and or Shapefile data
	- Converts point data to lines
	- converts lines to regularly sampled points
		04a_TrainingPt_TransectImport&Clean.Rmd
		04a_AA_TrainingPt_TransectImport&Clean.Rmd
		04b_TrainingPt_Summary.Rmd
			`r_avenza()`
			`make_lines()`
			`transect_sample()`

Field Data Points	- Spatial database of regular sample points - contains all empirical data for modelling - In the following steps this will be used to extract covariate data		

C. Chisholm

BC's Predictive Ecosystem Mapping Process

Part B: Modelling



Steps	Details	BEC DevExchange	pemgeneratr
Covariate Rasters	- Model raster resolution is set, and with an option for multiple resolutions		`multi_res()`
	- Covariates are generated remote sensing data. Typically these include derivatives from high resolution DTMs (e.g. from lidar). Additional options include: lidar forest canopy metrics, and/or other satelite data.		`create_covariates()`
	- Processing allows the stacking of multiple resolutions of the covariates		`cv_FineRes()`
Modelling Data	- Rasters are queried using the **Field Data Points**. - Provides a single data set that includes the empirical and remote sensed covariate data		`sample_covars()` PENDING
Split data: Test and Train	- 25% of the data is randomly selected from across all of the training data. - The emprical is stratified by class to ensure testing and training data contain the same ratio of empirical data.	05a_PEM_Build_Model.Rmd	

Create Model(s)	- Cross-fold validation is defined
	- model engine is defined: `ranger`
	- recipe is defined including:
	  - formula (e.g. `SiteSeries ~ .`)
	  - identify special columns such as:
	  - cLHS slices
	  - ID non-predictor data
	- workflow of above is defined
	- hyper-parameter tuning is defined
	- best tuning is selected
	- redefine model -- to include variable importance and
	- run last model -- includes analysis using split data (testing and training)
		05a_PEM_Build_Model.Rmd	`model_gen()` UPDATE NEEDED
			`inst/rmd/model_gen_tidy.Rmd`

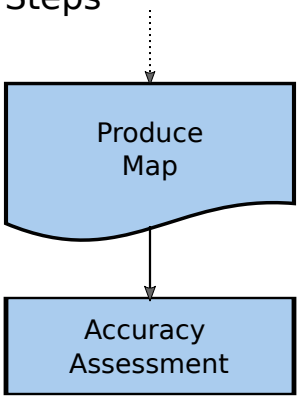
Model Report Model Object	- Exports: a) Report b) Model Object - used for predicting		

Note that the modelling framework started with the `mlr` package and is being converted to the `tidymodels` framework. This was done as `mlr` is no longer maintained, `mlr3` is not yet released, and the `tidymodels` systems is very well documented. The main processing engine, `ranger` (used to process random forest models), remains the same.

BC's Predictive Ecosystem Mapping Process

Part C: Mapping and Map Accuracy

Steps	Details	BEC DevExchange	pemgeneratr

	- Uses a Model Object and the Covariate Rasters to produce a map	06a_PEM_Predict_Map.Rmd	`predict_landscape()`
			helper function:
			`tile_index()`

	- Methods are still in development	05b_PEM_external_AA_ Assessment.Rmd 06e_Model_Assessment.Rmd 07a_PEM_map_compare.Rmd	`model_aa()` PENDING `map_compare()` PENDING

Part D: Additional Functionality

Data balancing *(Optional)*	- Training data -- model balancing: including options to downsample, smote, or combinations thereof.	04d_TrainingPt_balancing.Rmd	`data_smote()` PENDING `data_downsample()` PENDING
Uncertainty Mapping	- Entropy -	04d_TrainingPt_balancing.Rmd	`model_uncertainty()`
Predict New Area	- Extend model to a new area -	06a_PEM_predict_new_ area_map.RMD	Uses exiting functions
Map Cleaning	- Clean up the map raster -	07d_Despeckle_and_ Polygonize.R	`map_depeckle()` PENDING
Polygonize	- Converts the raster to polygons	07d_Despeckle_and_ Polygonize.R	`map_depeckle()` PENDING

C. Chisholm