Data Dictionary: Covariates by Domain

C.J. Tinant with ChatGPT 4o (20250504)

Contents

Introduction 1

Topography 1

Climate 4

Land Cover 6

Watershed Metrics 8

Introduction

This document summarizes the covariate metadata used in the regional skew estimation study. Variables are grouped by **Domain** and described in terms of their name, definition, unit, spatial scale, and analytical notes.

Topography

Table 1: Table A: Topography — Definition and Units

Short Name	Description	Unit
Station Altitude	Gage-level metadata representing [coordinate/altitude], used as a geospatial covariate for modeling spatial gradients.	meters
Mean Slope	Mean slope (from NED), aggregated to macroregional zones.	degrees
Median Slope	Median slope (from NED), aggregated to macroregional zones.	degrees
Altitude Zone	Mean elevation of gage locations, aggregated to macroregional zones.	meters
Mean Slope Distribution	Zonal mean of 10 m NED slope values, aggregated to Level II ecoregions.	degrees
Slope Skewness	Zonal skewness of 10 m NED slope values, aggregated to Level II ecoregions.	unitless (skewness)
Slope Variability (StDev)	Zonal standard deviation (SD) of slope values from 10 m NED data, aggregated to Level II ecoregions.	degrees
Slope Variability (IQR)	Zonal interquartile range (IQR) of 10 m NED slope values, aggregated to Level II ecoregions.	degrees
Surface Roughness	Zonal median of surface roughness, derived from 10 m NED, aggegated to Level III ecoregions.	m/m (slope ratio)
TWI Mean	Categorical version of the Topographic Wetness Index (TWI), created by binning continuous TWI values into ordinal classes (e.g., quartiles or hydrologically relevant thresholds). Simplifies terrain–moisture interactions for modeling and interpretability.	unitless index

Table 1: Table A: Topography — Definition and Units *(continued)*

Short Name	Description	Unit
TWI Modal	Zonal mode of 10 m NED Topographic Wetness Index (TWI) values aggegated to Level III ecoregions.	unitless index
TWI Class	Zonal assignment of ordinal NED Topographic Wetness Index (TWI) classes, derived from 10 m NED, aggegated to Level III ecoregions.	class label (ordinal ID)
Elevation Range	Zonal median elevation range (max - min elevation) from 10 m NED data, aggregated to NHD+ catchments.	meters
Aspect (Cosine)	Zonal median of the cosine of aspect, derived from 10 m NED data, aggregated to NHD+ catchments.	dimensionless [–1 to 1]
Aspect (Sine)	Zonal median of the sine of aspect, derived from 10 m NED data, aggregated to NHD+ catchments.	dimensionless [–1 to 1]
Curvature IQR	IQR of combined curvature types, derived from 10 m NED data, aggregated to NHD+ catchments.	unitless
Planform Curvature	Zonal median of planform (horizontal) curvature, derived from 10 m NED data, aggregated to NHD+ catchments.	unitless
Profile Curvature	Zonal median of profile (vertical) curvature, derived from 10 m NED data, aggregated to NHD+ catchments.	unitless
Watershed Elongation	Watershed elongation ratio, calculated from catchment geometry, aggregated to NHD+ catchments.	unitless ratio (m per m)
Watershed Circularity	Watershed circularity ratio, calculated from catchment geometry, aggregated to NHD+ catchments.	unitless ratio (sq-m per sq-m)

Table 2: Table B: Topography — Notes and Temporal Scale

Short Name	Notes	Scale
Station Altitude	Elevation of the gage site above sea level. Influences temperature, snowpack, and runoff timing	Station
Mean Slope	Describes average terrain steepness. Helps distinguish rugged uplands from flatter plains.	Macroregional
Median Slope	Reflects typical terrain slope. More robust to outliers than the mean; useful for evaluating runoff tendency.	Macroregional
Altitude Zone	Represents elevational regime. Influences climate, snow duration, and runoff seasonality.	Macroregional
Mean Slope Distribution	Aggregated mean slope at ecoregion scale. Affects stream power, flow velocity, and sediment dynamics.	Regional
Slope Skewness	Measures asymmetry in slope distribution. Helps identify dominant terrain types and slope transitions.	Regional
Slope Variability (StDev)	Quantifies slope variability including extremes. Suggests rugged or heterogeneous terrain.	Regional
Slope Variability (IQR)	Measures slope variability using IQR. Robust to outliers; reflects terrain complexity and infiltration dynamics.	Regional
Surface Roughness	Captures fine-scale surface texture. Influences ponding, microrelief, and infiltration pathways.	Subregional
TWI Mean	Indicates average terrain wetness potential. Combines slope and upstream drainage area. Lower classes indicate drier conditions; higher classes represent wetter zones with greater runoff accumulation potential.	Subregional
TWI Modal	Shows most frequent wetness class. Useful for identifying dominant moisture or runoff conditions.	Subregional

Table 2: Table B: Topography — Notes and Temporal Scale *(continued)*

Short Name	Notes	Scale
TWI Class	Ordinal wetness class derived from TWI. Simplifies terrain–moisture interactions; lower values = drier, higher = wetter zones.	Subregional
Elevation Range	Captures elevation spread within catchment. Linked to slope energy, erosion risk, and runoff response.	Local
Aspect (Cosine)	Quantifies N–S slope orientation. North-facing slopes (values > 0) retain snow longer and reduce evapotranspiration.	Local
Aspect (Sine)	Quantifies E–W slope orientation. East-facing slopes (values > 0) warm earlier, affecting snowmelt and phenology.	Local
Curvature IQR	Measures variation in terrain curvature. High values imply complex flow paths and erosion susceptibility.	Local
Planform Curvature	Identifies slope convergence/divergence. Negative = concave (channels), positive = convex (ridges).	Local
Profile Curvature	Represents slope change along flow lines. Helps distinguish depositional vs. erosional terrain zones.	Local
Watershed Elongation	Describes catchment shape. More elongated basins have faster runoff and flashier peaks.	Local
Watershed Circularity	Captures catchment compactness. Higher values suggest faster, more concentrated runoff response.	Local

Climate

Table 3: Table A: Climate — Definition and Units

Short Name	Description	Unit
Longitude	Gage-level metadata representing [coordinate/altitude], used as a geospatial covariate for modeling spatial gradients.	decimal degrees
Latitude	Gage-level metadata representing [coordinate/altitude], used as a geospatial covariate for modeling spatial gradients.	decimal degrees
Climate Zone	Dominant Köppen–Geiger climate class, aggregated to macroregional zones.	categorical code
PHZM Zone Count	Count of unique USDA Plant Hardiness Zones, aggregated to macroregional zones.	count
PHZM Zone (Dominant)	Dominant USDA Plant Hardiness Zone, aggregated to macroregional zones.	categorical code
Annual Temp (Median)	Zonal median of annual mean temperature from PRISM 1991–2020 climatological normals, aggregated to Level II ecoregions.	degrees C
Annual Precip (Median)	Zonal median of total annual precipitation from PRISM 1991–2020 climatological normals, aggregated to Level II ecoregions.	mm
Pct Annual Precip May–Aug (Mean)	Zonal mean of the percentage of annual precipitation falling May–August from PRISM 1991–2020 climatological normals, aggregated to Level II ecoregions.	proportion [0–1]
Pct Annual Precip May–Aug (Median)	Zonal median of the percentage of annual precipitation falling May–August from PRISM 1991–2020 climatological normals, aggregated to Level II ecoregions.	proportion [0–1]
Seasonal Precip StDev	Standard deviation of seasonal precip totals from PRISM 1991–2020 normals aggegated to Level III ecoregions.	mm
Seasonal Precip IQR	IQR (25th–75th percentile) of seasonal precip totals from PRISM 1991–2020 normals aggegated to Level III ecoregions.	mm
Fall Precip (Median)	Sum of zonal medians of fall monthly precip totals (Sep–Nov) from PRISM 1991–2020 normals aggegated to Level III ecoregions.	mm
Winter Precip (Median)	Sum of zonal medians of winter monthly precip (Dec–Feb) from PRISM 1991–2020 normals aggegated to Level III ecoregions.	mm
Summer Precip (Median)	Sum of zonal medians of spring monthly precip (Mar–May) from PRISM 1991–2020 normals aggegated to Level III ecoregions.	mm
Spring Precip (Median)	Sum of zonal medians of summer monthly precip (Jun–Aug) from PRISM 1991–2020 normals aggegated to Level III ecoregions.	mm
Freeze–Thaw Days	Zonal count of freeze–thaw transition days from PRISM daily data (1991–2020), aggegated to NHD+ catchments.	days
Precip Intensity (95th-percentile)	Zonal median of 95th percentile of daily precip on wet days from PRISM daily normals (1991–2020), aggegated to NHD+ catchments.	mm
Wet Day Frequency (Seasonal)	Zonal median number of wet days (geq 0.1 mm/day) per season from PRISM daily normals (1991–2020), aggregated to NHD+ catchments.	days

Table 4: Table B: Climate — Notes and Temporal Scale

Short Name	Notes	Scale
Longitude	Represents site location in degrees east or west. Useful for detecting climatic or physiographic gradients (moisture availability, storm type).	Station
Latitude	Represents site location in degrees north. Commonly used to capture latitudinal patterns in climate and hydrology (e.g., temperature, seasonality).	Station
Climate Zone	Provides broad-scale climate context (e.g., humid continental vs. semi-arid). Useful for regional stratification.	Macroregional
PHZM Zone Count	Reflects cold hardiness diversity; transitional zones may influence vegetation, snowmelt timing, and freeze–thaw cycles.	Macroregional
PHZM Zone (Dominant)	Captures dominant cold tolerance regime based on minimum winter temperatures. Relevant for snowmelt timing and freeze–thaw sensitivity.	Macroregional
Annual Temp (Median)	Governs evapotranspiration, freeze–thaw, and snowpack formation. Represents the baseline thermal regime.	Regional
Annual Precip (Median)	Represents typical annual moisture supply; important for soil moisture availability, runoff potential, and vegetation productivity.	Regional
Pct Annual Precip May–Aug (Mean)	Characterizes warm-season water input. Useful for identifying storm-driven systems tied to convective seasonality.	Regional
Pct Annual Precip May–Aug (Median)	Captures the typical proportion of precipitation falling in the warm season—important for Great Plains hydrology.	Regional
Seasonal Precip StDev	Captures variability across seasons. Sensitive to outliers—useful for detecting flashiness or hydrologic irregularity.	Subregional
Seasonal Precip IQR	Characterizes intra-annual precipitation variability; robust to outliers.	Subregional
Fall Precip (Median)	Captures fall precipitation; relevant to baseflow reset and transition to frontal storm systems.	Subregional
Winter Precip (Median)	Indicates snowpack formation potential and winter water storage.	Subregional
Summer Precip (Median)	Captures convective storm seasonality (flood-prone regions).	Subregional
Spring Precip (Median)	Key springtime recharge and flood generation season; relevant to melt-transition systems and early crop needs.	Subregional
Freeze–Thaw Days	Captures thermally transitional zones; relevant for soil structure, infiltration disruption, and freeze–thaw runoff events.	Local
Precip Intensity (95th-percentile)	Captures storm severity at the upper tail. Correlated with flash flood potential and runoff generation efficiency.	Local
Wet Day Frequency (Seasonal)	Describes storm frequency and seasonal spacing. Interacts with intensity to shape hydrograph form and flood likelihood.	Local

Land Cover

Table 5: Table A: Land Cover — Definition and Units

Short Name	Description	Unit
Cropland Fraction	Fraction of cropland (NLCD), aggregated to macroregional zones.	proportion [0–1]
Forest Fraction	Fraction of forested land (NLCD), aggregated to macroregional zones.	proportion [0–1]
Grassland Fraction	Fraction of grassland (NLCD), aggregated to macroregional zones.	proportion [0–1]
Urban Fraction	Fraction of urban land (NLCD), aggregated to macroregional zones.	proportion [0–1]
Land Cover % (MODIS)	Zonal median of percent cover of dominant land class (cropland, forest, etc.) from MODIS 2016 land cover, aggegated to Level III ecoregions.	proportion [0–1]
Land Use Diversity (MODIS)	Normalized Shannon-Weiner Diversity Index of MODIS land classes aggegated to Level III ecoregions.	unitless index (Shannon-Weiner)
Land Cover % (NLCD)	Zonal median percent cover by dominant NLCD class aggegated to NHD+ catchments.	proportion [0–1]
Land Use Diversity (NLCD)	Normalized Shannon–Weiner Diversity Index of NLCD land cover classes, aggregated to NHD+ catchments.	unitless index (Shannon-Weiner)
NDVI Amplitude	Zonal median of annual NDVI amplitude for 2016 MODIS data, aggregated to Level II ecoregions.	unitless [–1 to 1]
NDVI IQR	Zonal median of the interquartile range (IQR) of NDVI values for 2016 MODIS data, aggregated to Level II ecoregions.	unitless [–1 to 1]
Growing Season Length	Zonal median of the length of the 2016 growing season (NDVI-based), aggregated to Level II ecoregions.	days
Peak NDVI	Zonal median of the seasonal peak NDVI for 2016 MODIS data, aggregated to Level II ecoregions.	unitless [–1 to 1]

Table 6: Table B: Land Cover — Notes and Temporal Scale

Short Name	Notes	Scale
Cropland Fraction	Indicates anthropogenic land use in agricultural regions. Associated with increased runoff, reduced infiltration, and higher ET demand.	Macroregional
Forest Fraction	Forest cover enhances interception and storage. Typically reduces peak flows and lowers skew.	Macroregional
Grassland Fraction	Moderates evapotranspiration and infiltration rates; especially important in mixed rangeland systems.	Macroregional
Urban Fraction	Urban development increases imperviousness; strongly associated with peak flow generation and positive flood skew.	Macroregional
Land Cover % (MODIS)	Summarizes dominant land cover class by proportion within a zone. Helps interpret complex or fragmented landscapes by reducing pixel-level detail to a single representative class.	Subregional
Land Use Diversity (MODIS)	Reflects land use heterogeneity. Higher values indicate fragmented or mixed land cover, which may affect infiltration and runoff patterns.	Subregional
Land Cover % (NLCD)	Shows dominant land cover proportion at the catchment scale. Useful for linking terrain features with hydrologic behavior.	Local
Land Use Diversity (NLCD)	Measures land cover diversity within a catchment. Supports modeling of transitions or mosaic land use patterns.	Local

Table 6: Table B: Land Cover — Notes and Temporal Scale *(continued)*

Short Name	Notes	Scale
NDVI Amplitude	Captures seasonal variation in greenness. High amplitude reflects strong vegetation response to climate fluctuations.	Regional
NDVI IQR	Describes central spread of seasonal greenness. Robust to outliers and well-suited to fragmented or transitional landscapes.	Regional
Growing Season Length	Indicates length of vegetative activity. Influences evapotranspiration rates and baseflow timing.	Regional
Peak NDVI	Represents seasonal greenness maximum. Serves as a proxy for peak vegetation productivity.	Regional

Watershed Metrics

Table 7: Table A: Watershed Metrics — Definition and Units

Short Name	Description	Unit
Clay Fraction	Zonal modal fraction of surface soil (0–5 cm) composed of clay particles from SSURGO2, aggregated to Level II ecoregions.	proportion [0–1]
Silt Fraction	Zonal modal fraction of surface soil (0–5 cm) composed of silt particles from SSURGO2, aggregated to Level II ecoregions.	proportion [0–1]
Sand Fraction	Zonal modal fraction of surface soil (0–5 cm) composed of sand particles from SSURGO2, aggregated to Level II ecoregions.	proportion [0–1]
Soil Permeability	Zonal median of STATSGO2 permeability values aggegated to Level III ecoregions.	mm/h
Soil Runoff Class	Zonal modal soil hydrologic group (A–D) from STATSGO2 aggegated to Level III ecoregions.	class A–D (ordinal)
Watershed Area	Gage-level metadata representing contributing watershed area.	watershed area (sq-km)
Median Stream Order	Zonal median of Strahler stream order, aggregated to Level II ecoregions.	Strahler order
Max Stream Order	Zonal maximum stream order in the network, aggregated to Level II ecoregions.	Strahler order
Flow Accumulation	Zonal median of 10 m NED flow accumulation (upslope contributing area) aggegated to Level III ecoregions.	contributing area (sq-km)
Stream Density	Zonal median of stream length per catchment area, derived from NHD+ data.	meters per square meter
Mean Flow Length	Zonal median of average flow path length, derived from NHD+ data.	meters
Stream Slope	Zonal median of stream slope (elevation drop per unit flow length), derived from NHD+ data.	slope ratio (m/m)
Relief Ratio	Relief ratio (elevation range divided by stream length) from 10 m NED data, aggregated to NHD+ catchments.	m/m

Table 8: Table B: Watershed Metrics — Notes and Temporal Scale

Short Name	Notes	Scale
Clay Fraction	Indicates surface soil texture. Higher clay content reduces infiltration and increases runoff potential.	Regional
Silt Fraction	Indicates intermediate soil texture. Balances water retention and infiltration capacity.	Regional
Sand Fraction	Represents infiltration rate of surface soils. Low values indicate higher runoff and steeper hydrographs.	Regional
Soil Permeability	Represents infiltration rate of surface soils. Low values indicate higher runoff and steeper hydrographs.	Subregional
Soil Runoff Class	USDA hydrologic soil group (A–D). Classifies soils by infiltration capacity and storm response.	Subregional
Watershed Area	Commonly used as a spatial predictor in hydrologic modeling.	Station
Median Stream Order	Represents typical drainage network complexity. Based on Strahler order, indicating average stream hierarchy within the watershed.	Regional
Max Stream Order	Indicates the largest stream order within the catchment. Useful for identifying drainage basin scale and extent.	Regional
Flow Accumulation	Estimates upslope contributing area. Higher values suggest greater potential for water accumulation and flooding.	Subregional

Table 8: Table B: Watershed Metrics — Notes and Temporal Scale *(continued)*

Short Name	Notes	Scale
Stream Density	Ratio of total stream length to catchment area. Higher density reflects greater drainage dissection and shorter runoff travel times.	Local
Mean Flow Length	Average flow path length to outlet. Longer paths typically indicate delayed runoff response.	Local
Stream Slope	Elevation drop per unit stream length. Influences runoff velocity and hydrograph shape.	Local
Relief Ratio	Ratio of watershed relief to main channel length. High values indicate steep basins and higher hydrologic energy.	Local