There are 5 data files in csv format (comma separated value).

Some quick notes:

- The model runs from 800 to 1350 in annual time steps. So the year variable is from 800 to 1350.
- The landscape is defined on a map of 80 × 120 cells (Figure 1 from "Understanding Artificial Anasazi" paper). Each cell represents a 100m × 100m space.

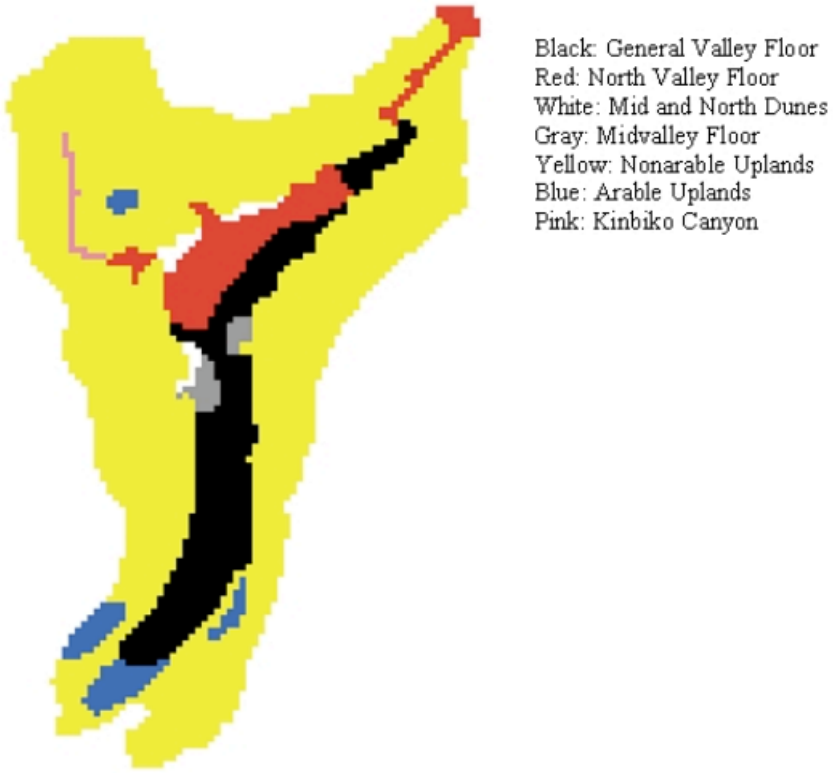

Figure 1. Different zones of land cover

File 1/5: map.csv

Contains information about the zones along with other attributes like colours so you can re-create Figure 1.

Column	Type	Possible values	Note
x	int	0-79	x coordinate in the simulation
y	Int	0-119	y coordinate in the simulation
color	string	"white" "yellow" "pink" "blue" "red" "black" "gray"	match with Figure 1
zone	string	"Empty" "Natural" "Kinbiko" "Uplands" "North" "General" "North Dunes" "Mid Dunes" "Mid"	
maize.zone	string	"Empty" "No_Yield" "Yield_1" "Yield_3" "Yield_2" "Sand_dune"	Coressponding to the yield levels for different zones (Table 3 in "Understanding Artificial Anasazi" paper) • "Yield_1" is "North and Mid Valley, Kinbiko Canyon" column. • "Yield_2" is "General Valley" column. • "Yield_3" is "Arable Uplands" column. • "Sand_dune" is "Dunes" column.

Table 3: The yield levels for the different values of PDSI

	Zones			
PDSI	North and Mid Valley, Kinbiko Canyon	General Valley	Arable Uplands	Dunes
(−∞,−3]	617	514	411	642
(−3,−1]	719	599	479	749
(−1,1)	821	684	547	855
[1,3)	988	824	659	1030
[3,∞)	1153	961	769	1201

File 2/5: settlements.csv

Historical settlements data (real world settlements to compare to our simulated settlements).

You may ignore this for now. We plan to post-process this dataset and provide a time series of the real world settlements.

Column	Type	Note
SARG number		an ID number relating to data collected by the "Southwestern Anthropological Research Group" (unused by the simulation, left for reference)
meter north	int	real world location measurements
meter east	int	real world location measurements
start date	int	historical start date
end date	int	historical end date
median date	int	historical median date
settlement type		extra data (unused)
settlement size		extra data (unused)
description		extra data (unused)
room count		extra data (unused)
elevation		extra data (unused)
baseline households		baseline number of households
number of households		"current" number of households
x	int	x coordinate in the simulation
y	int	y coordinate in the simulation

File 3/5: water.csv

Location of water sources

Column	Type	Note
id number	int	an ID number (unused)
meter north	int	real world location measurements
meter east	int	real world location measurements
type	int	1 = conditional on zones and year. 2 = always a water source. 3 = water source when the current year is between start and end date.
start date	int	historical start date
end date	int	historical end date
x	int	x coordinate in the simulation
y	int	y coordinate in the simulation

Pseudo code for water sources:

```
//NOTE: 1 is TRUE, 0 is FALSE

if ((year >= 280 and year < 360) or (year >= 800 and year < 930) or (year >= 1300
and year < 1450))

       existStreams = 1

else

       existStreams = 0

if (((year >= 420) and (year < 560)) or ((year >= 630) and (year < 680)) or
((year >= 980) and (year < 1120)) or ((year >= 1180) and (year < 1230)))

       existAlluvium = 1

else

       existAlluvium = 0


for each location {

       water-source = 0

       if (existAlluvium == 1) and ((zone == "General") or (zone == "North") or
(zone == "Mid") or (zone == "Kinbiko")) {

              water-source = 1

       }

       if (existStreams == 1) and (zone == "Kinbiko") {

         water-source = 1

       }

}
```

```
for these locations: (location 72 114) (location 70 113) (location 69 112)
(location 68 111) (location 67 110) (location 66 109) (location 65 108) (location
65 107)) {

       water-source = 1

}


// water-type, water-x, water-y are the column names, read from file

for each water-points-from-file {

       if water-type = 2 {

              (location with water-x water-y).water-source = 1 //the location with
matching x,y coor has water source

       }

       if water-type = 3 {

              if year >= start-date and year <= end-date {

                     (location with water-x water-y).water-source = 1

              }

       }

}
```

File 4/5: pdsi.csv

The adjust PDSI (Palmer Drought Severity Index) for each zone over time (corresponding to the PDSI column of Table 3).

There are 7 columns: the year columns (800 to 1350) and the 6 columns for 6 zones.

File 5/5: hydro.csv

Water availability for each zones (value from 0 to 10). This is the level of water (different to "whether there is water" from water.csv which is either 0 or 1).

There are 7 columns: the year columns (800 to 1350) and the 6 columns for 6 zones.

Note: After finding a new farm location, the household will settle to a nearby location that is not a farm and has hydro<=0.