current processes for matching Native Forest Law enrolled properties to CIRN cadastre

Albert Garcia

1 matching enrolled Native Forest Law properties to CIRN rural properties

1.1 data

Native Forest Law dataset contains the following variables:

```
#merging project property, owner, and
projects_df <- proyecto_df %>%
  inner_join(predio_df, by = c("rptpro_id", "rptpro_numero_ingreso")) %>%
  left_join(propietario_df, by = "rptpre_id") %>%
  inner_join(coordinadas_predio_df, by = "rptpre_id")

colnames(projects_df)
```

```
##  [1] "rptpro_id"
##  [2] "rptpro_numero_ingreso.x"
##  [3] "rptpro_puntaje"
##  [4] "rptpro_superficie"
##  [5] "rptpro_tipo_concurso"
##  [6] "rptpro_objetivo_manejo"
##  [7] "rptpro_tipo_postulacion"
##  [8] "rptpro_tipo_presentacion"
##  [9] "rptpro_fecha_presentacion"
## [10] "rptpro_ano"
## [11] "rptpro_literal"
## [12] "rptpro_region"
## [13] "rptpro_id_concurso"
## [14] "rptpro_numero_region"
## [15] "rptpro_tipo_presenta"
## [16] "rptpro_presentado_por"
## [17] "rptpro_monto_total"
## [18] "rptpro_monto_ordenacion"
## [19] "rptpro_monto_actividades"
## [20] "rptpro_aporte"
## [21] "rptpro_nombre_concurso"
## [22] "rptpro_superficie_predios_postulante"
## [23] "rptpre_id"
## [24] "rptpre_nombre"
## [25] "rptpre_rol"
## [26] "rptpre_superficie_bonificada"
## [27] "rptpre_monto_ordenacion"
```

```
## [28] "rptpre_superficie_predial"
## [29] "rptpre_comuna"
## [30] "rptpre_provincia"
## [31] "rptpre_region"
## [32] "rptpre_monto_total"
## [33] "rptpre_monto_actividades"
## [34] "rptpre_ordenacion"
## [35] "rptpre_aporte"
## [36] "rptprop_id"
## [37] "rptprop_nombre"
## [38] "rptprop_apellido_paterno"
## [39] "rptprop_apellido_materno"
## [40] "rptprop_tipo"
## [41] "rptprop_razon_social"
## [42] "rptprop_ciudad"
## [43] "rptprop_sexo"
## [44] "rptprop_etnia"
## [45] "rptpro_numero_ingreso.y"
## [46] "rptub_id"
## [47] "rptub_referencia"
## [48] "rptub_datum"
## [49] "rptub_huso"
## [50] "rptub_norte"
## [51] "rptub_este"
```

Some important variables for matching to CIRN data:

rptpre_rol - property rol id

rptprop_nombre - property name

rptprop_nombre - owner name

rptub_norte, rptub_este - property location northing and easting respectively

rptub_datum - corresponding property datum

rptub_huso - corresponding utm zone

We also have stand data with the following variables, but I will not use them at this point:

```
colnames(rodal_df)
```

```
##  [1] "rptro_id"                  "rptpre_id"
##  [3] "rptro_numero"              "rptro_superficie"
##  [5] "rptro_tipo_forestal"       "rptro_categoria_conservacion"
##  [7] "rptro_especie"             "rptro_datum"
##  [9] "rptro_huso"                "rptro_norte"
## [11] "rptro_este"                "rptro_monto"
## [13] "rptro_aporte"              "rptro_numero_bonificacion"
## [15] "rptro_numero_plan"         "rptro_bosque_quemado"
```

CIRN data contain the following relevant variables to match with Natove Forest Law properties:

```r
colnames(prop_rural.sf)
```

```
##  [1] "ROL"         "PROPIETARI" "NOM_PREDIO" "SUPERFICIE" "RIEGO1_HA"
##  [6] "RIEGO2_HA"   "RIEGO3_HA"  "RIEGO4_HA"  "SECAN1_HA"  "SECAN2_HA"
## [11] "SECAN3_HA"   "SECAN4_HA"  "SECAN5_HA"  "SECAN6_HA"  "SECAN7_HA"
## [16] "SECAN8_HA"   "DESCCOMU"   "region"     "AP_PATERBN" "AP_MATERBN"
## [21] "NOMBRESBN"   "NUMERO"     "NOMBRE"     "geometry"   "my_prop_id"
```

Some important variables to consider are:

ROL - property rol id

PROPIETARI - owner name

NOM_PREDIO - property name

1.2 Native Forest Law property coordinates

```r
# turn into spatial object
projects_df.sf <- st_as_sf(projects_df, coords = c( "rptub_este", "rptub_norte"), na.fail = FA

#table(projects_df.sf$rptub_datum)
# trying to clean entries into intended datum

#if datum contains, 84, wgs, or WGS, I change the name to WGS84
projects_df.sf$rptub_datum <- gsub(".*(84|wgs|WGS).*", "WGS84", projects_df.sf$rptub_datum)
#similar for SAD69 and PSAD56
projects_df.sf$rptub_datum <- gsub(".*(56|PSAD).*", "PSAD56", projects_df.sf$rptub_datum)
projects_df.sf$rptub_datum <- gsub(".*69*.", "SAD69", projects_df.sf$rptub_datum)
#table(projects_df.sf$rptub_datum)

#new dataframe for projects with WGS84 and utm zone 18 south
projects_WGS19 <- projects_df.sf %>%
  filter(rptub_huso == 19 & rptub_datum == "WGS84")%>%
  st_set_crs("+proj=utm +zone=19 +south +datum=WGS84") %>%  #set crs
  st_transform(st_crs(prop_clean.sf))   #change crs to same as CIRN data

projects_WGS84 <- projects_df.sf %>%
  filter(rptub_huso != 19 & rptub_datum == "WGS84")%>%
  st_set_crs("+proj=utm +zone=18 +south +datum=WGS84") %>%
  st_transform(st_crs(prop_clean.sf)) %>%
  rbind(projects_WGS19)
```

Of the 15239 projects with property coordinates, the cleaned dataframe had 14554 in WGS84 datum, 376 in SAD69 and 288 in PSAD56. For now, I just focus on the WGS84 projects. Of those WGS84 projects, 13825 had an assigned UTM zone as either 18 or 19. I perform the join using these 13825 projects.

1.3 joining by location

```
proj_enrolled <- prop_clean.sf %>%
  st_join(projects_WGS84, join = st_intersects) %>%
    drop_na(rptpro_id) # drop properties without a Native Forest Law project id
```

We can visually examine some of the properties below to address the validity of the join:

Table 1: spatially joined dataframe

	ROL	rptpre_rol	NOM_PREDIO	rptpre_nombre
164427	1391-9	1391-42	AGUA FRIA	AGUA FRIA LOTE B
251916	1324-1	1324-1	AGUA LUNA-PEDREGOSO	AGUA LUNA
259617	331-10	333-12; 333-12	EL AROMO	AGUA SANTA
217043	309-18	309-18	HONDURAS	AGUAS CLARAS
217043.1	309-18	309-18	HONDURAS	AGUAS CLARAS
165018	1386-23	1386-24	EL LLEUQUE	Aguas Coloradas
165018.1	1386-23	1386-24	EL LLEUQUE	Aguas Coloradas
165018.2	1386-23	1386-24	EL LLEUQUE	Aguas Coloradas
128799	111-14	111-20	HJ EL MANIGUAL	AGUAS DEL PASO MALO
157397.1	164-33	164-268	SECTOR BULARCO SECTOR A COMUNA DE RANQUI	Aguas negras parcela B2 lote 3 Bularco
157397	164-33	164-268	SECTOR BULARCO SECTOR A COMUNA DE RANQUI	AGUAS NEGRAS PC B2 LOTE 3 BULARCO
308209.1	526	1220-59	NA	Ahuenco Parcela 19
176162	539-34	539-33	STA AGUSTINA PTE	Alamo y Coihue
176162.1	539-34	539-33	STA AGUSTINA PTE	Alamo y Coihue
268027.2	112-16	112-32; 112-33	EL HUAPE	Alaska

Notice the first entry. The rol id does not match, but the property name is similar (Agua Fria vs. Ague Fria Lote B).

The number of unique Native Forest Law rol ids is :

```
## [1] 4431
```

Note also that, of the 13825 properties used in the join, 3592 did not fall in regions contained in our current CIRN study area (regions 5-10, 13). Here are the Native Forest Law projects by region:

```
##
##    1    2    3    4    5    6    7    8    9   10   11   12   13   14   16
##   31    4   33   11  319  923 1818 1456 3547 1963  867  243  207 1649  754
```

We see that the main excluded regions with a significant number of Native Forest Law projects are 11, 12, 14, and 16.

1.4 matching by rol id

We can also match by rol id between the CIRN and Native Forest Law datasets

```
#here, I just see how many rols match between the CIRN data and native forest law data

#projects_df is the native forest law projects
#prop_rural.sf is the CIRN data with the rols as simple features object
rol_df <- projects_df %>%
  rename(ROL = rptpre_rol) %>% #renaming rols to be the same in both datasets
  inner_join(prop_rural.sf, by = "ROL") #merge by rol
```

```r
# just creating a df to display only a couple of choice variables
rol_display <- rol_df %>%
  select(ROL, NOM_PREDIO, rptpre_nombre)
rol_display$geometry <- NULL

kable(rol_display[1:20,],
      "latex",
      booktabs =T,
      caption = "merged on rol dataframe"
      ) %>%
  kable_styling(latex_options = "hold_position", font_size = 6)
```

Table 2: merged on rol dataframe

ROL	NOM_PREDIO	rptpre_nombre
160-15	EL CARMEN	LOS MANZANOS
160-15	NA	LOS MANZANOS
160-15	BARANDICA HIJUELA 4	LOS MANZANOS
160-15	EL LITRE	LOS MANZANOS
160-15	LA HUACHA	LOS MANZANOS
160-15	PREDIO SAN AGUSTIN SECTOR RELBUN	LOS MANZANOS
160-15	PEUMO PICHILO S/N CARAMPANGUE	LOS MANZANOS
160-15	UITENTO	LOS MANZANOS
160-15	EL DAO AGUANTAO HJ 48	LOS MANZANOS
304-121	EL ESCUDO PARCELA N 18	PARCELA N°18
304-121	HIJUELA 7 HEULIO COM CATREFOL	PARCELA N°18
304-121	EL ESCUDO PARCELA N 18	PARCELA N°18
304-121	HIJUELA 7 HEULIO COM CATREFOL	PARCELA N°18
3290-9	EL MIRADOR PC 21 VILCUN TEMUCO	PARCELA N°21
3290-9	EL MIRADOR PC 21 VILCUN TEMUCO	PARCELA N°21
3290-9	EL MIRADOR PC 21 VILCUN TEMUCO	PARCELA N°21
3290-9	EL MIRADOR PC 21 VILCUN TEMUCO	PARCELA N°21
135-10	CAM A CHINCOLC	QUETROLELFU
135-10	LA MONTANA COM	QUETROLELFU
135-10	PARCELA N 1 PATRIA NUEVA EX FUNDO MOSTAZ	QUETROLELFU

```r
length(unique(rol_df$ROL)) #find number of unique rol numbers found
```

```
## [1] 4263
```

Visually examining the exposes an interesting aspect of the CIRN data, which is that one rol id has multiple different property names associated to it. I'm unsure what the rol id corresponds to exactly, but found this interesting.

Using either method, we match about half of the potential native forest law properties (roughly 4500 out of 10233). We can look into more sophisticated ways to match. (maybe some combination of location, rol id, and part of property name)

1.5 Spatially joining using the provided stand coordinates

The number of unique Native Forest Law rol ids using the stand coordinates rather than the property is:

```
## [1] 4059
```

Table 3: spatially joined dataframe based on stand location

	ROL	rptpre_rol	NOM_PREDIO	rptpre_nombre
360803.16	137-43	137-43/44	CERROS ABRANTES	Abrantes Maderera
360803.17	137-43	137-43/44	CERROS ABRANTES	Abrantes Maderera
360803.18	137-43	137-43/44	CERROS ABRANTES	Abrantes Maderera
360803.19	137-43	137-43/44	CERROS ABRANTES	Abrantes Maderera
360835	137-44	137-43/44	EL MANZANITO LOTE B	Abrantes Maderera
360835.1	137-44	137-43/44	EL MANZANITO LOTE B	Abrantes Maderera
360835.2	137-44	137-43/44	EL MANZANITO LOTE B	Abrantes Maderera
360835.3	137-44	137-43/44	EL MANZANITO LOTE B	Abrantes Maderera
307129	752	1230-30	NA	Aconcagua
322616	1	141-70	NA	AGONI BAJO - MARIA AND
178770.7	205-17	205-31	SAN AGUSTIN PTE	Agricola Los Temos
178770.8	205-17	205-31	SAN AGUSTIN PTE	Agricola Los Temos
268815	242-5	180-31; 180-33; 180-77; 181-9; 242-1; 242-2;242-3; 242-4;242-5;242-28;242-29	CARMEN DE TRAFUN	Agricola Trafun Spa
268866	180-31	180-31; 180-33; 180-77; 181-9; 242-1; 242-2;242-3; 242-4;242-5;242-28;242-29	LAS HIJUELAS	Agricola Trafun Spa
268966	242-29	180-31; 180-33; 180-77; 181-9; 242-1; 242-2;242-3; 242-4;242-5;242-28;242-29	LAS HIJUELAS	Agricola Trafun Spa
273825.3	75-25	75-25	COIQUE HJ 4	Agroganadera
273825.4	75-25	75-25	COIQUE HJ 4	Agroganadera
273825.5	75-25	75-25	COIQUE HJ 4	Agroganadera
273825.6	75-25	75-25	COIQUE HJ 4	Agroganadera
273825.7	75-25	75-25	COIQUE HJ 4	Agroganadera
273825.8	75-25	75-25	COIQUE HJ 4	Agroganadera
151234.2	VP-E	220-192	NA	Agua de la Perdíz
163198	SR	220-192	NA	Agua de la Perdíz
164427	1391-9	1391-42	AGUA FRIA	AGUA FRIA LOTE B
164427.1	1391-9	1391-42	AGUA FRIA	AGUA FRIA LOTE B
251916	1324-1	1324-1	AGUA LUNA-PEDREGOSO	AGUA LUNA
251916.1	1324-1	1324-1	AGUA LUNA-PEDREGOSO	AGUA LUNA
251916.2	1324-1	1324-1	AGUA LUNA-PEDREGOSO	AGUA LUNA
165641	267-33	267 - 73	LOS NAVIOS	Agua Santa
165641.1	267-33	267 - 73	LOS NAVIOS	Agua Santa