

02_comparacion_obs

Pao

June 10, 2019

SATWND

La idea es comparar los datos de viento del GOES que se usaron durante relampago con los que están disponibles en el prepbufr (SATWND). Cómo tengo prepbufr para ventanas de 1 hora centradas voy a comparar las obs de las 18 UTC del 20181120 (o sea las obs entre las 17:30 y las 18:30) con las observaciones en ese período.

```
obs_pb <- fread('../obs_PREPBUFR/20182011_18_prepbufr.csv',
                na.strings = "10000000000.000")
colnames(obs_pb) <- c("obs.id", "time", "station", "lon", "lat", "dhr", "sub.id", "station.elev", "p", "u", "v")

obs_pb[, time := ymd_h(time)]
obs_pb[, time.obs := time + as.period(as.duration(days(x=1))*(dhr/24))]

obs_pb <- obs_pb[obs.id == "SATWND"] %>%
  .[lon %between% c(290, 310) & lat %between% c(-40, -20)]

#Le saco variables

knitr::kable(obs_pb[, .N, by = .(sub.id)])
```

sub.id	N
253	153
243	89
254	161

```
obs_pb <- obs_pb[, list(lon, lat, p, u, v, time.obs)] %>%
  melt(., measure.vars = c("u", "v"), variable.name = "obs.id", value.name = "obs") %>%
  .[obs.id == "u", obs.id := "2819"] %>%
  .[obs.id == "v", obs.id := "2820"] %>%
  .[, source := "prepbufr"]
```

Estos son los tipos de observación dentro de “SATWND” y corresponden a:

Report Type	PREPBUFR Message Type	Description	Ind.
241	SATWND	INDIA IR AND VISIBLE CLOUD DRIFT U, V (INSAT-2E)	+
242	SATWND	JMA IR AND VISIBLE CLOUD DRIFT U, V AT LEVELS BELOW 850 MB (GMS-5)	A
243	SATWND	EUMETSAT IR AND VISIBLE CLOUD DRIFT U, V AT LEVELS BELOW 850 MB (METEOSAT-5, METEOSAT-7)	A
244	—	currently not used	n/a
245	SATWND	NESDIS IR CLOUD DRIFT U, V (GOES-8, GOES-10)	A
246	SATWND	NESDIS IMAGER WATER VAPOR CLOUD U, V AT CLOUD TOP (GOES-8, GOES-10)	A

Report Type	PREPBUFR Message Type	Description	Ind.
247	SATWND	NESDIS IMAGER WATER VAPOR CLOUD U, V - DEEP LAYER (GOES-8, GOES-10)	*
248	SATWND	NESDIS SOUNDER WATER VAPOR CLOUD U, V AT CLOUD TOP (GOES-8, GOES-10)	*
249	SATWND	NESDIS SOUNDER WATER VAPOR CLOUD U, V - DEEP LAYER (GOES-8, GOES-10)	*
250	SATWND	JMA WATER VAPOR CLOUD U, V (GMS-5)	+
251	SATWND	NESDIS VISIBLE CLOUD DRIFT U, V (GOES-8, GOES-10)	+
252	SATWND	JMA IR AND VISIBLE CLOUD DRIFT U, V AT LEVELS ABOVE 850 MB (GMS-5)	A
253	SATWND	EUMETSAT IR AND VISIBLE CLOUD DRIFT U, V AT LEVELS ABOVE 850 MB (METEOSAT-5, METEOSAT-7)	A
254	SATWND	EUMETSAT WATER VAPOR CLOUD U, V (METEOSAT-5, METEOSAT-7)	+
255	SATWND	NESDIS PICTURE TRIPLET CLOUD U, V (GOES-8, GOES-10)	A
256	SATWND	INDIA WATER VAPOR CLOUD U, V (INSAT-2E)	*

Así que entiendo que GOES-16 no estaría llegando a los prepbufs...

Pero, encontré la siguiente tabla que contendría los códigos actualizados y un ejemplo donde se explica que significan los caracteres que hay en el lugar de la estación. Para mis prepbuf como no se de donde vienen voy a dejar ese dato como "XXXXXX".

De todo lo anterior concluyo que el **el código para GOES-16 es 270**

```
obs_rra <- read.obs('../obs_RRA/20181120_18/*0.dat', keep.obs = c(2819,2820)) %>%
  rbind(read.obs('../obs_RRA/20181120_19/*0.dat', keep.obs = c(2819,2820)))

#Filtro solo lo que necesito (time range, dominio, tipo de u,v)
obs_rra <- obs_rra[time %between% c(as_datetime("2018-11-20 17:30:00"), as_datetime("2018-11-20 18:30:00")),
  .[sub.id == 4] %>% #Asumo que sub.id = 4 es para los datos del goes porque son muchísimos y al graficar se ven bien
  .[lon %between% c(290, 310) & lat %between% c(-40, -20)]

obs_rra <- obs_rra[, list(lon, lat, elev, time, obs.id, obs)] %>%
  .[, source := "rra"]

setnames(obs_rra, old = c("time", "elev"), new = c("time.obs", "p"))

# Ojo! la elevación está en hectopascasles!
```

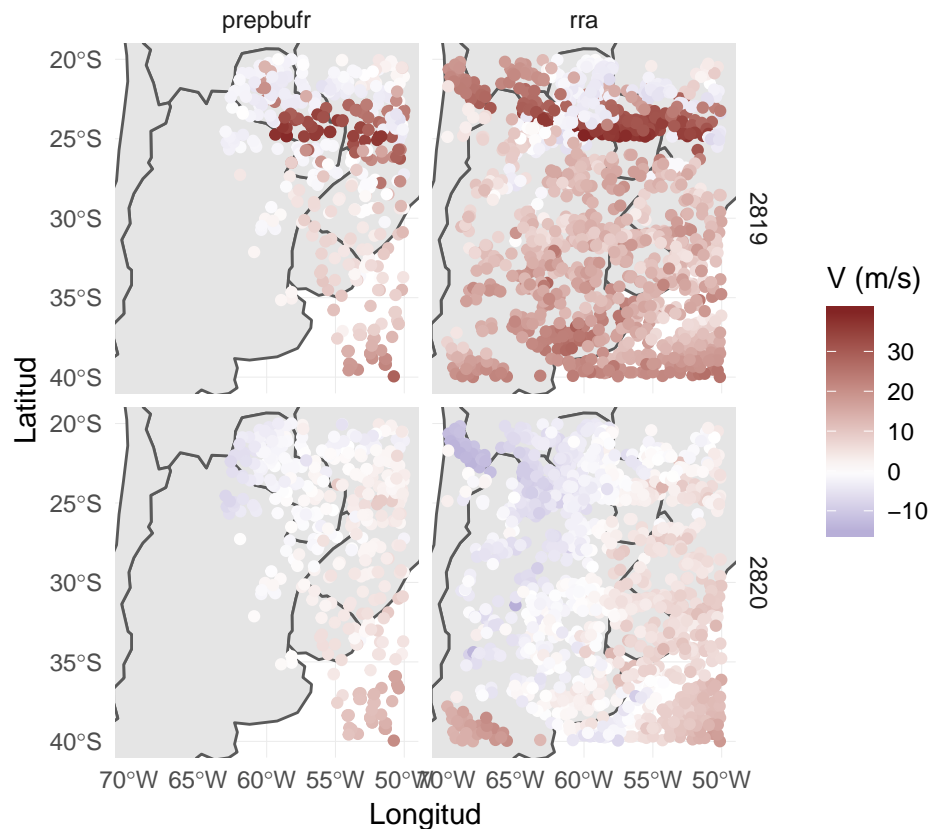
Tengo dos datatable con el mismo formato, vamos a ver que pinta tienen.

```
obs <- rbind(obs_pb, obs_rra)

map <- rnaturalearth::ne_countries(continent = 'south america', scale = 'small', returnclass = 'sf')

ggplot(obs, aes(ConvertLongitude(lon), lat)) +
  geom_sf(data = map, inherit.aes = FALSE) +
  coord_sf(xlim = c(-70, -50), ylim = c(-40, -20)) +
  geom_point(aes(color = obs)) +
  scale_color_divergent(name = "V (m/s)") +
```

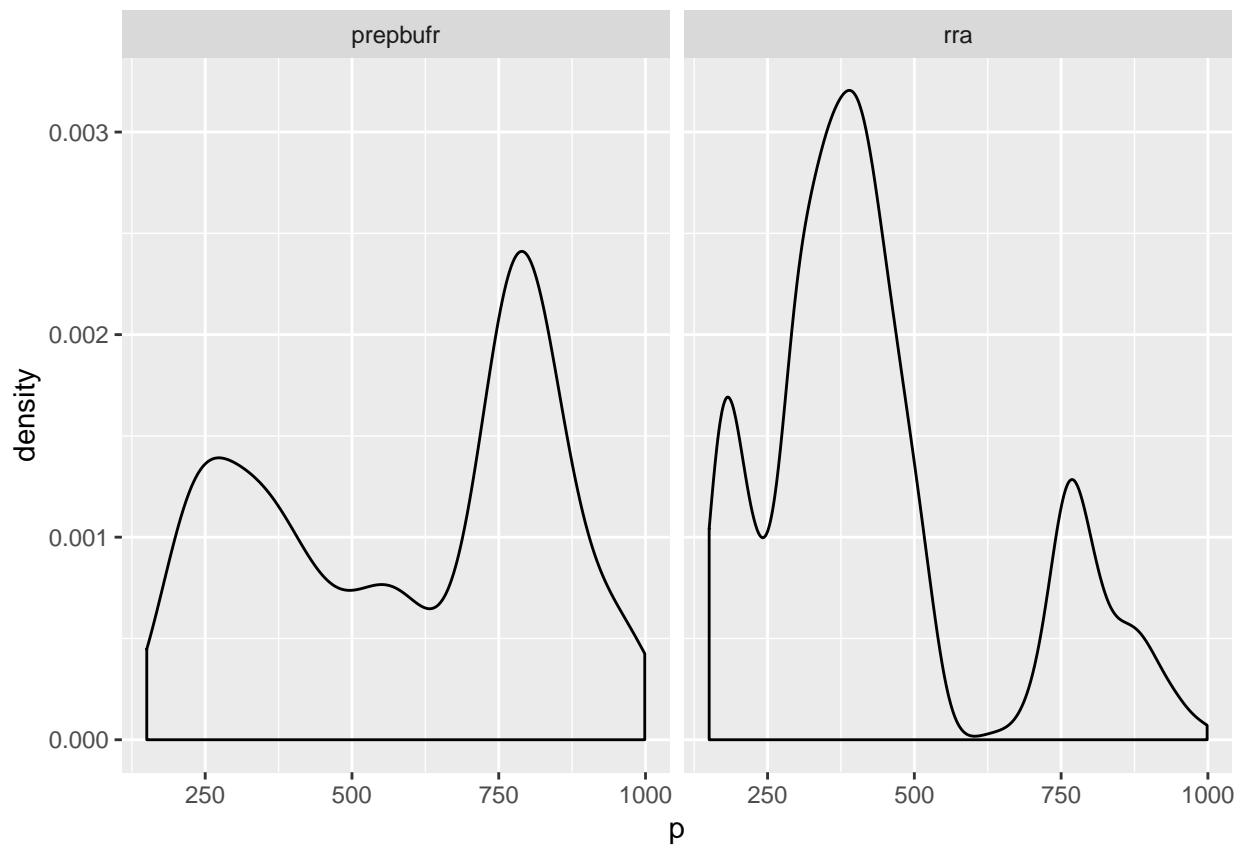
```
scale_x_continuous(name = "Longitud") +
scale_y_continuous(name = "Latitud") +
facet_grid(obs.id~source) +
theme_minimal()
```



En la figura se comparan las observaciones de los dos dataset. Los datos usados para el RRA son más y tienen mayor cobertura respecto de los que vienen en el prepbufr. Según las referencias de prepbufr los datos son de Eumetsat, no me queda claro si esto incluye o no a los vientos del GOES. Lo interesante es que el patrón se corresponde entre los dos data sets.

Cómo estamos trabajando con un campo 3D, es interesante ver la distribución en niveles (ojo, con las unidades). Ambos sets tienen un poco entre 750 y 800 hPa. En niveles más altos se ven claramente 2 máximos en ~400 hPa y en ~100 hPa para los datos ingestados por el RRA. Pero en los datos de prepbufr solo se ve un máximo en 250 hPa con mucha dispersión.

```
ggplot(obs, aes(p)) +
  geom_density() +
  facet_wrap(~source)
```



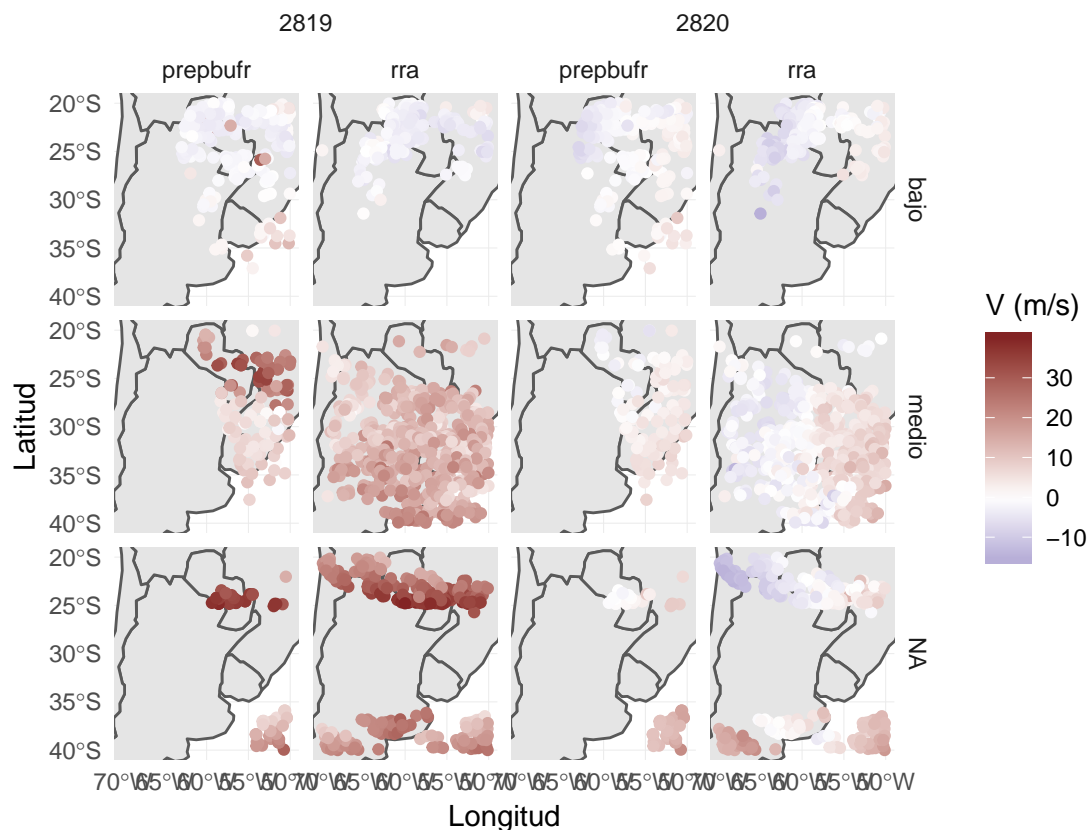
Vamos a intentar convertir la variable *p* (continua) en *capaz* (factores) para ver la distribución espacial por niveles.

Definimos:

- Nivel bajo: 1000 - 600 hPa.
- Nivel medio: 600 - 300 hPa.
- Nivel alto: menor a 300 hPa.

```
obs[p %between% c(600, 1000), lev := "bajo"] %>%
  .[p %between% c(300, 600), lev := "medio"] %>%
  .[p %between% c(300, 100), lev := "alto"]

obs[] %>%
  ggplot(aes(ConvertLongitude(lon), lat)) +
  geom_sf(data = map, inherit.aes = FALSE) +
  coord_sf(xlim = c(-70, -50), ylim = c(-40, -20)) +
  geom_point(aes(color = obs)) +
  scale_color_divergent(name = "V (m/s)") +
  scale_x_continuous(name = "Longitud") +
  scale_y_continuous(name = "Latitud") +
  facet_nested(lev~obs.id + source) +
  theme_minimal()
```



Si bien el patrón general se mantiene, es más fácil ver algunas diferencias. En niveles bajos el prepbufr tiene algunos datos anormalmente altos, da la sensación de ser errores porque no aparecen en el otro set de datos y están en una región con viento zonal casi nulo. En niveles medios las diferencias son más importantes: La región de viento zonal fuerte sobre el este de Paraguay que muestra el prepbufr no se observa en el otro dataset. Se me ocurre que puede haber un problema con la estimación de los niveles y que esos puntos en realidad corresponden a niveles altos.

AIRCFT

Hagamos el mismo análisis pero para datos de aviones. En el prepbufr estos datos están codificados con “AIRCFT” y en los datos del RRA corresponden al subtipo 3.

Report Type	PREPBUFR Message Type	Description	Ind.
230	AIRCFT	AIREP AND PIREP AIRCRAFT U, V	A
231	AIRCFT	ASDAR AIRCRAFT U, V	A

```
# prepbufr
obs_pb <- fread('../obs_PREPBUFR/20182011_18_prepbufr.csv',
  na.strings = "10000000000.000")
colnames(obs_pb) <- c("obs.id", "time", "station", "lon", "lat", "dhr", "sub.id", "station.elev", "p", "v")

obs_pb[, time := ymd_h(time)]
obs_pb[, time.obs := time + as.period(as.duration(days(x=1))*(dhr/24))]

obs_pb <- obs_pb[obs.id == "AIRCFT"] %>%
```

```

.[lon %between% c(290, 310) & lat %between% c(-40, -20)]

obs_pb <- obs_pb[, list(lon, lat, p, u, v, time.obs)] %>%
  melt(., measure.vars = c("u", "v"), variable.name = "obs.id", value.name = "obs") %>%
  .[obs.id == "u", obs.id := "2819"] %>%
  .[obs.id == "v", obs.id := "2820"] %>%
  .[, source := "prepbufr"]

# RRA
obs_rra <- read.obs('../obs_RRA/20181120_18/*0.dat', keep.obs = c(2819,2820)) %>%
  rbind(read.obs('../obs_RRA/20181120_19/*0.dat', keep.obs = c(2819,2820)))

#Filtro solo lo que necesito (time range, dominio, tipo de u,v)
obs_rra <- obs_rra[time %between% c(as_datetime("2018-11-20 17:30:00"), as_datetime("2018-11-20 18:30:00")), ]
.[sub.id == 3] %>%
.[lon %between% c(290, 310) & lat %between% c(-40, -20)]

obs_rra <- obs_rra[, list(lon, lat, elev, time, obs.id, obs)] %>%
  .[, source := "rra"]

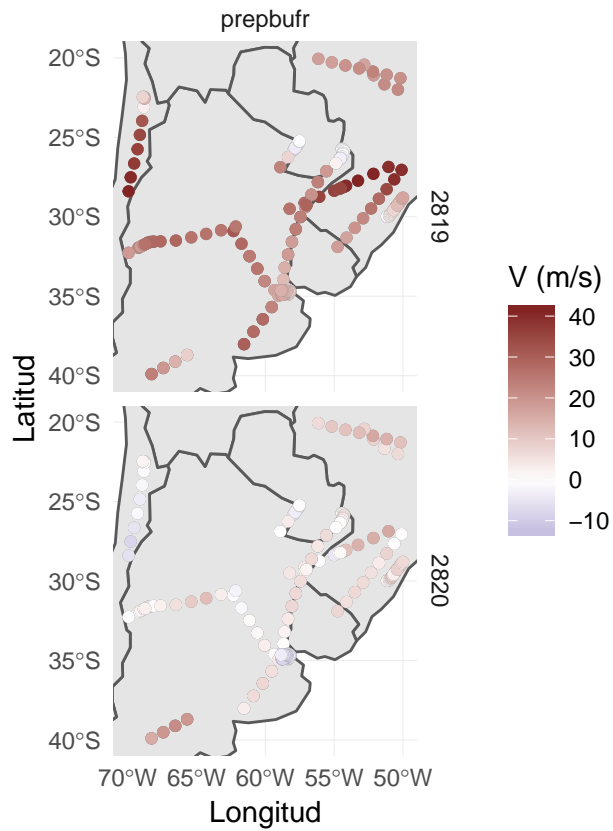
setnames(obs_rra, old = c("time", "elev"), new = c("time.obs", "p"))

# La unión
obs <- rbind(obs_pb, obs_rra)

map <- rnaturalearth::ne_countries(continent = 'south america', scale = 'small', returnclass = 'sf')

ggplot(obs, aes(ConvertLongitude(lon), lat)) +
  geom_sf(data = map, inherit.aes = FALSE) +
  coord_sf(xlim = c(-70, -50), ylim = c(-40, -20)) +
  geom_point(aes(color = obs)) +
  scale_color_divergent(name = "V (m/s)") +
  scale_x_continuous(name = "Longitud") +
  scale_y_continuous(name = "Latitud") +
  facet_grid(obs.id~source) +
  theme_minimal()

```



Bueno, buenas noticias! Al parecer los prebufrs tienen más observaciones de aviones que el RRA y por suerte las que están en el RRA también están en el prebufr.

AIRS

En principio los prebufrs no tienen datos de AIRS, pero deberían estar en un bufr aparte ¿serán los mismo datos?

En los dato susados en RRA el subtipo para AIRS es 21. Pero no tengo ninguno en este tiempo.