

Résumé semaine du 20 février 2023

gbbeaudet

24/02/2023

ANALYSES EFFECTUÉES CETTE SEMAINE:

Pas énormément d'analyses graphiques cette semaine.

1. Compilation d'un nouveau jeu de donnée avec des identifiants unique pour chaque évènement de chaque BV, avec leur sévérité combinée et leur durée.
2. Identification d'erreurs et problèmes dans les données de l'atlas.
3. Récupération des données des tronçons hors Québec: Pas encore incluses dans le jeu de donnée du point 1.

1. Compilation du nouveau jeu de données Tout vas bien ici de manière générale. Par contre, identification d'évènements de longueurs irréalistes qui viennent insérer beaucoup d'erreurs (Plus d'info au point 2). À l'exception de ces évènements, les autres semblent bien correspondrent entre eux:

Figure 1.1 Overlap des évènements de sécheresses du BVs 0520 - Rivière Champlain

```
gg_vistime(events.df %>% filter(BV == "0502"),
  col.start = "start", col.end = "end", col.group = "id",
  col.event = "duration", col.color = "event.id")
```

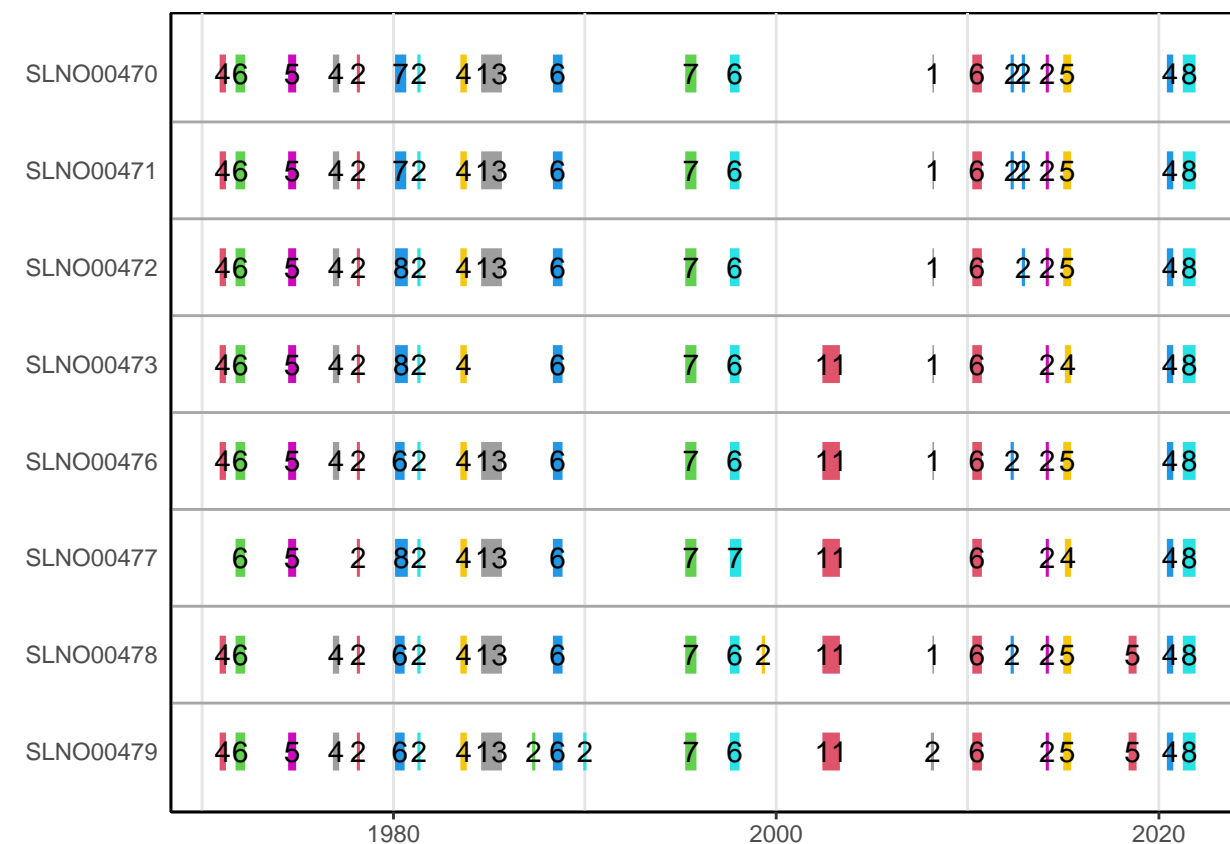


Figure 1.2 Overlap des évènements de sécheresses du BV 0220 - Rivière Rimouski

```
gg_vistime(events.df %>% filter(BV == "0220"),
  col.start = "start", col.end = "end", col.group = "id",
  col.event = "duration", col.color = "event.id")
```

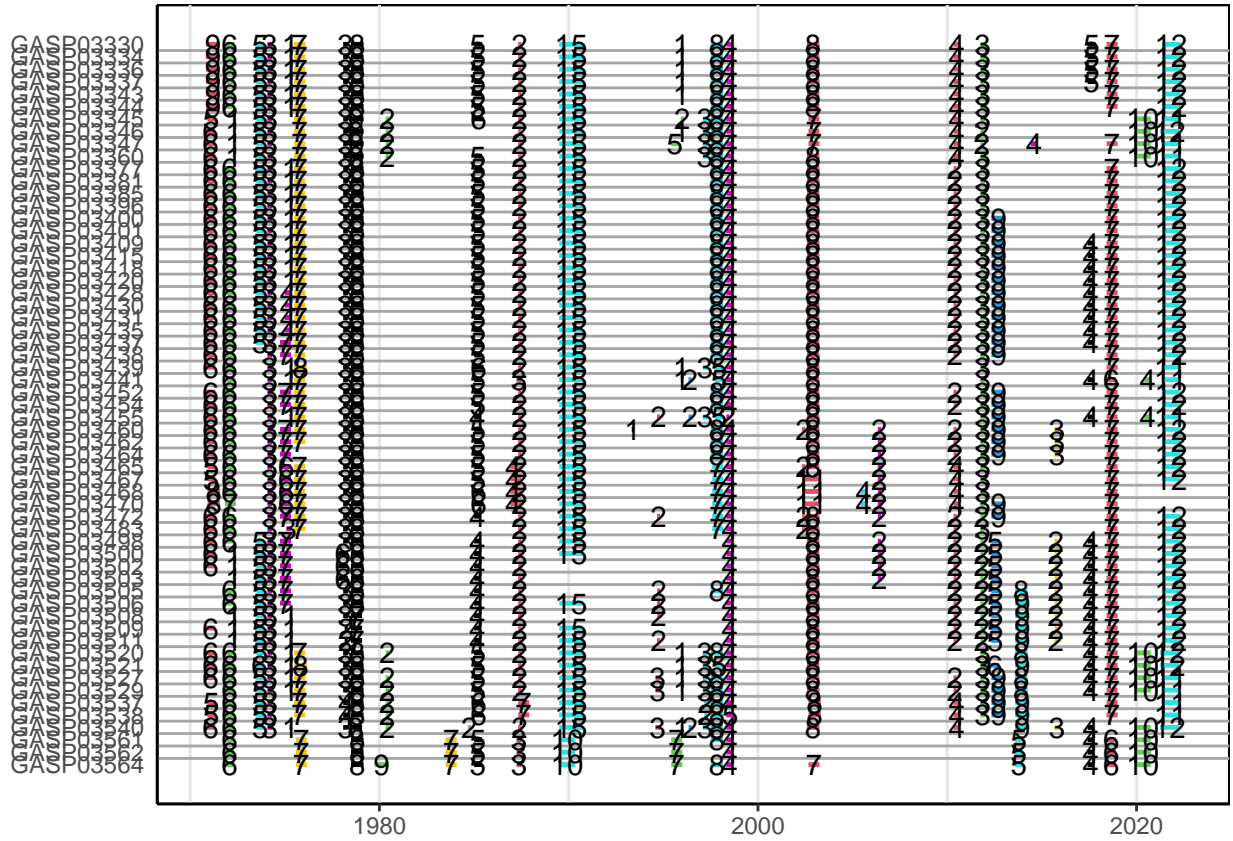


Figure 1.3 Overlap des événements de sécheresses du BV 0103 - Rivière du Grand Pabos

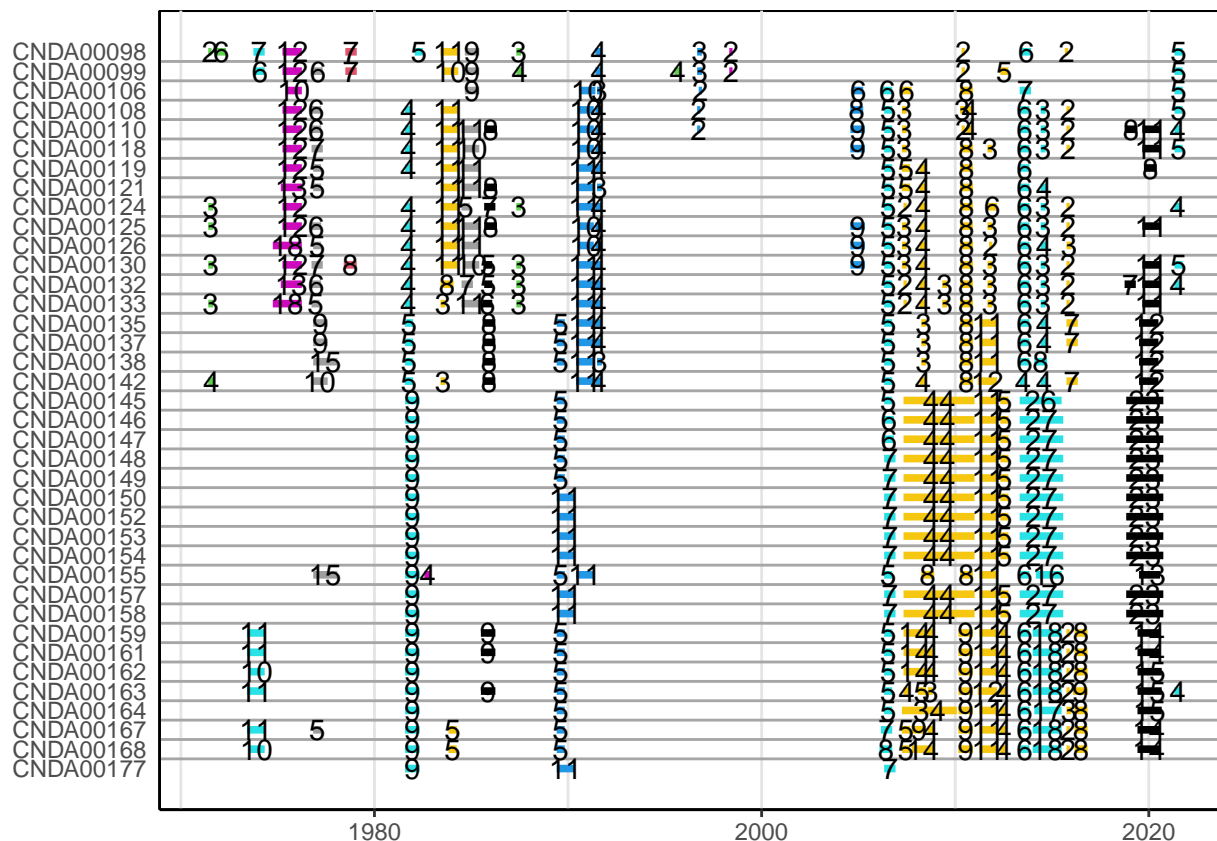
```
gg_vistime(events.df %>% filter(BV == "0103"),
  col.start = "start", col.end = "end", col.group = "id",
  col.event = "duration", col.color = "event.id")
```

| | | | | | | | | | |
|-----------|---|----|------|----|----|-------|------|---|-------|
| GASP01895 | 6 | 34 | 26 | 11 | 65 | 48 | | 2 | 104 |
| GASP01897 | 6 | 34 | 26 | 11 | 65 | 48 | | 2 | 104 |
| GASP01898 | 6 | 34 | 26 | 11 | 65 | 48 | | 2 | 104 |
| GASP01900 | 6 | 34 | 26 | 11 | 65 | 48 | | 2 | 104 |
| GASP01902 | 6 | 34 | 26 | 11 | 65 | 48 | | 2 | 104 |
| GASP01903 | 9 | 34 | 26 | 11 | 62 | 10 49 | | 2 | 4104 |
| GASP01904 | 6 | 34 | 26 | 11 | 65 | 48 | | 2 | 104 |
| GASP01905 | 9 | 34 | 26 | 11 | 62 | 49 | | 2 | 104 |
| GASP01906 | 6 | 34 | 26 | 10 | 65 | 48 | | 2 | 104 |
| GASP01907 | 6 | 4 | 26 | 10 | 65 | 48 | 32 | 2 | 104 |
| GASP01908 | 6 | 4 | 26 | 11 | 65 | 48 | | 2 | 104 |
| GASP01909 | 6 | 4 | 26 | 11 | 67 | 48 | | 2 | 6 104 |
| GASP01911 | 6 | 4 | 7 26 | 11 | 67 | 47 | | 2 | 6 104 |
| GASP01915 | 6 | 34 | 26 | 8 | 3 | 48 | 35 6 | 2 | 6 104 |
| GASP01917 | 6 | 34 | 26 | 8 | 63 | 48 | 35 6 | 2 | 6 104 |
| GASP01919 | 6 | 34 | 26 | 8 | 3 | 48 | 35 6 | 2 | 6 104 |
| GASP01921 | 6 | 34 | 26 | 8 | 3 | 48 | 35 6 | 2 | 6 104 |
| GASP01924 | 7 | 34 | 26 | 8 | 5 | 3 48 | | 2 | 104 |

1980 2000 2020

Figure 2. Exemple des problèmes causés par un évènement de longueur irréaliste

```
gg_vistime((events.df %>% filter (BV == "0710"))[1:500,],
  col.start = "start", col.end = "end", col.group = "id",
  col.event = "duration", col.color = "event.id")
```



Prochaine Étape: Régler ces erreurs puis faire correspondre les identifiants des évènements entre les tous les BVs à l'échelle du Québec.

2. Identification d'erreurs et problèmes dans les données de l'atlas. Les évènements de très longues durées (~ 30 mois et plus) viennent principalement de 3 régions hydrographiques:

- 04 - Outaouais et Montréal
- 06 - Saguenay
- 07 - Côte-Nord

La raison de ces longueurs irréalistes est une déviation prononcée apr rapport à l'abscisse 0 pour des périodes couvrant plusieurs années.

D'abord, il n'y a pas énormément de BV qui montre cette variation, et lorsqu'elle est présente, c'est généralement la majorité des tronçons. Donc il n'y a pas vraiment de distinction avec la superficie de drainage du tronçons. Il ne semble pas avoir de distinction avec la superficie du BV non plus.

La période problématique n'est pas toujours la même Par contre, elle est généralement la même au sein d'une région hydrographique entière. Il y a un pattern distinctif dans tous les BV de la région 07, la 06, et la 04. 07 est la côte-nord dont les données sont incertaines selon le CEHQ, 04 contient la rivière des Outaouais qui est diviser un peu étrangement, et 06 est le Saguenay qui n'a normalement rien de vraiment problématique. **Après discussion, on réalise la présence de barrages importants dans des BVs qualifiés de "non influencé par un barrage", notamment la rivière au Outardes (#0710), sur laquelle 4 importants barrages de Hydro-Québec se trouvent.** Voir les figures suivantes. Il reste à

voir si les autres déviations ont aussi des barrages dans leur BV respectif, et sinon de comprendre d'où elles proviennent.

Figure 3. Série temporelle de SSI du BV #0710 - Rivière aux outardes.

```
bv <- 710

n.sect <- mean( (df %>% filter(BV == bv))$n.sect)
d.area.BV <- round(max( (df %>% filter(BV == bv))$d.area),0)
BV.id <- max((df %>% filter(BV == bv))$BV)

ggplot(df %>% filter(BV == bv))+
  geom_line(aes( x = timestamp, y = SSI.gam, group = id, col = d.area),alpha = .5)+
  theme_classic() + scale_color_continuous(type = "viridis")+
  labs( title = paste("BV:",BV.id," ", "n.sect," tronçons, ",d.area.BV," km2", sep = ""))
```

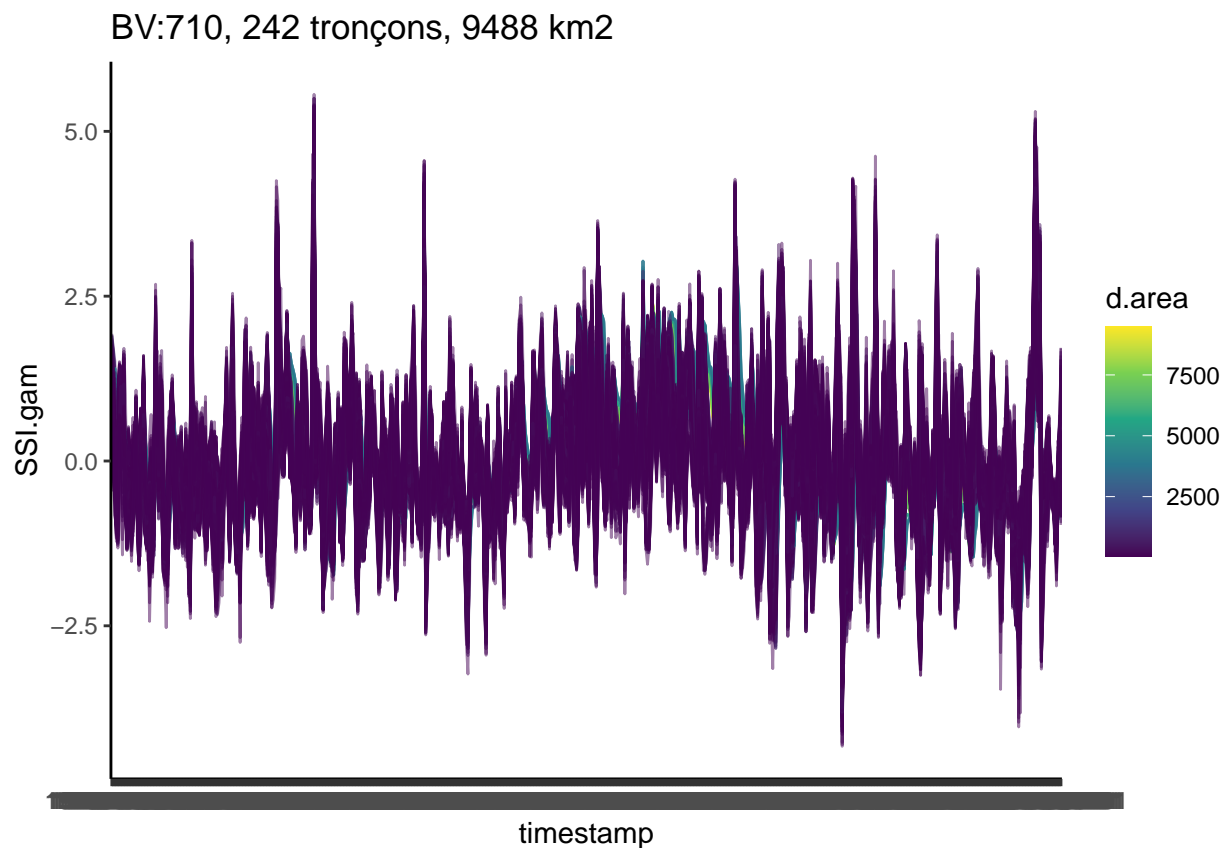


Figure 4. Série temporelle de SSI du BV #0710 - Rivière aux outardes, tronçon “CNDA00178”.

```
plot.ts((df %>% filter(id == "CNDA00178"))$SSI.gam)
```

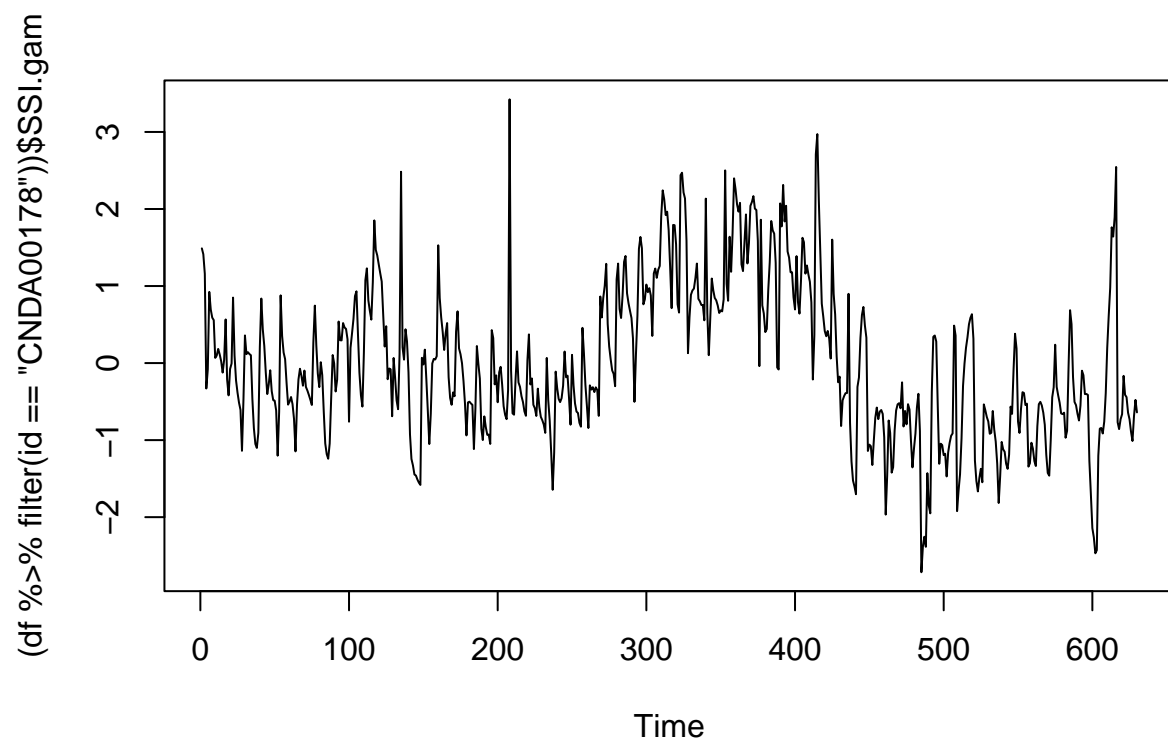
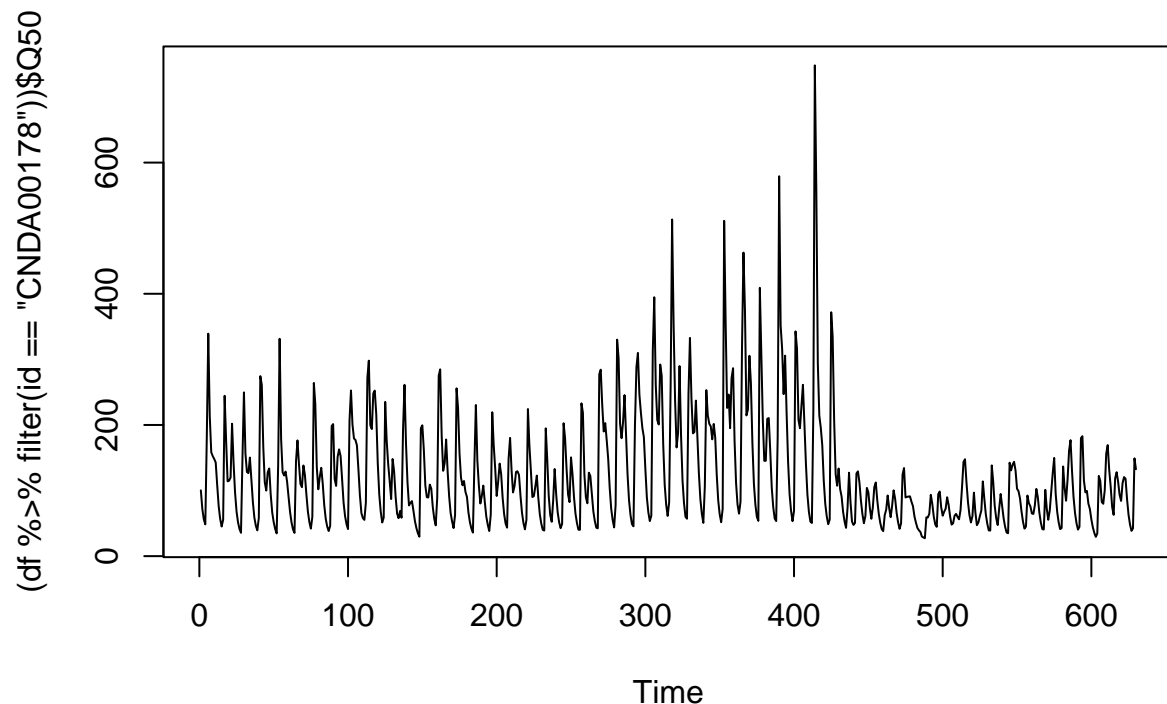


Figure 5. Série temporelle des débits du BV #0710 - Rivière aux outardes

```
plot.ts((df %>% filter(id == "CNDA00178"))$Q50)
```



Une autre possibilité est la suivante, tirée d'un courriel envoyé à Audrey concernant la série temporelle ci-haut:

Ces patterns causent problèmes parce que ça donne des évènements de longueurs impossibles et ça vient booster la variabilité des évènements dans le BV. Ça cache aussi plusieurs plus petits évènements qui ressortirait si on recentrait les grandes déviations alentour de zéro. Ça cause aussi problème quand vient le temps d'overlapper les évènements entre les tronçons pour leurs donner un identifiant, parce que le gros évènement va venir englober plusieurs petits et tous les mettre sous le même identifiant. Avec notre discussion de tout à l'heure je réalise quelque chose. Cette série temporelle ne respecte pas du tout les seuils de 3 SSI de -1.5, etc. On voit dans la série qu'il y a au moins 10 mois en dessous de -1.5 à la fin, et sûrement pas loin de 5 ans au dessus de 1. Alors là on a deux possibilités. Soit les débits de ces tronçons (possiblement ceux qui ont une trop grande surface de drainage) ont des distributions peu communes, donc le fitting d'une distribution théorique est mauvais et la transposition sur une loi normale aussi, ou bien la distribution que l'on a sélectionnée ne correspond pas très bien à ces tronçons, menant au même résultat. Ceci se confirme assez bien avec les exemples que j'ai regardé.

Courriel de Gabriel à Audrey, le 23 février.

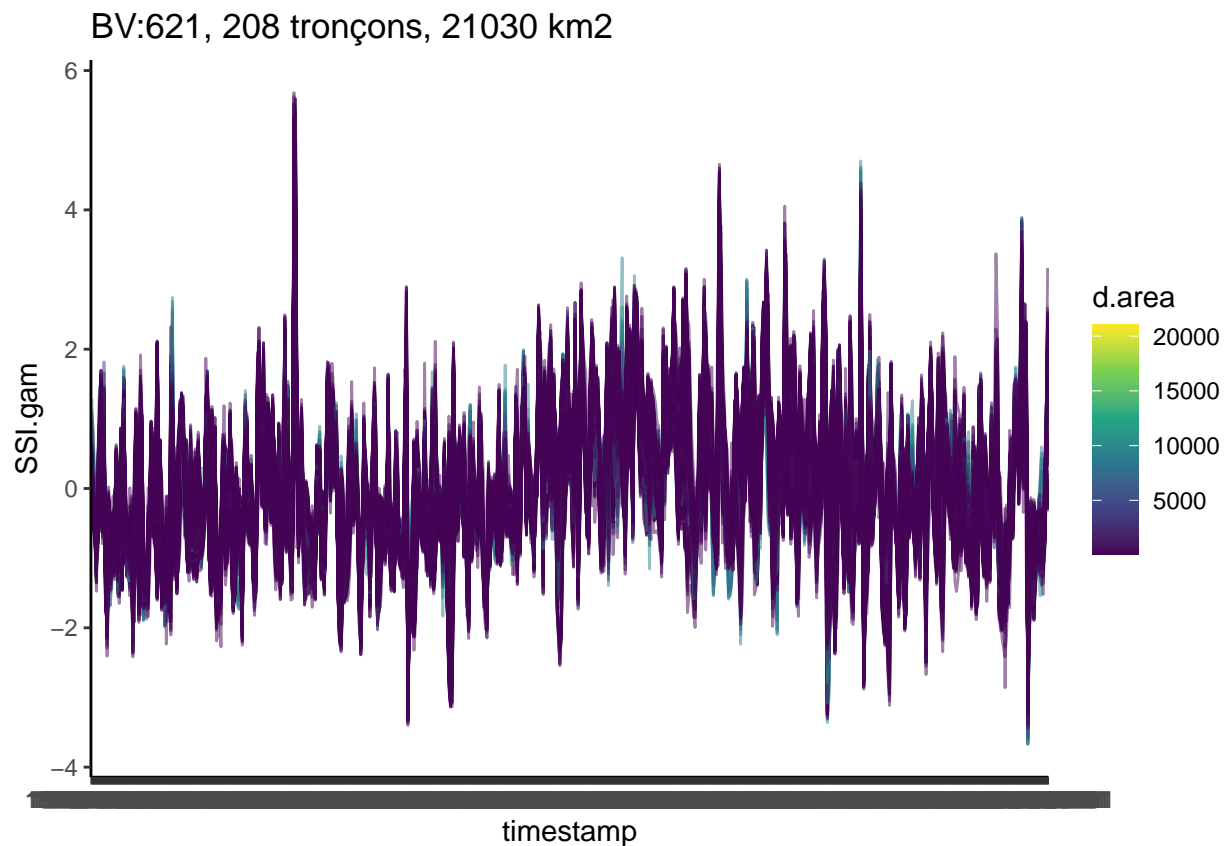
Une note finale: Même dans les grands BVs, sérieusement, les séries des tronçons se ressemblent beaucoup. Par exemple, ci-dessous le # 0621, avec 208 tronçons. Il y a beaucoup de variabilité au niveau de la magnitude des SSIs, mais au niveau temporelle les évènements tombent bien ensemble. Je pense que donc que la prédiction que les évènements sévères sont généralisés au BV et que les plus faibles sont plus dispersés fait du sens (Hypothèse 1, voir document google docs: "2023-02-16_hypothese-prediction.docx").

Figure 6. Série temporelle des débits du BV #0621 - Rivière Mistassini

```
bv <- 621

n.sect <- mean( (df %>% filter(BV == bv))$n.sect)
d.area.BV <- round(max( (df %>% filter(BV == bv))$d.area),0)
BV.id <- max((df %>% filter(BV == bv))$BV)

ggplot(df %>% filter(BV == bv))+
  geom_line(aes( x = timestamp, y = SSI.gam, group = id, col = d.area),alpha = .5)+
  theme_classic() + scale_color_continuous(type = "viridis")+
  labs( title = paste("BV:",BV.id," ", "n.sect," tronçons, ",d.area.BV," km2", sep = ""))
```



Les séries temporelles des SSI et des débits de *tous* les BVs sont disponibles sous forme de PDF sur le portable personnel de Gabriel ([~/R_proj/Drought_events/](#))