

EBOV empirical sampling proportions

Louis du Plessis

October 31, 2018

Contents

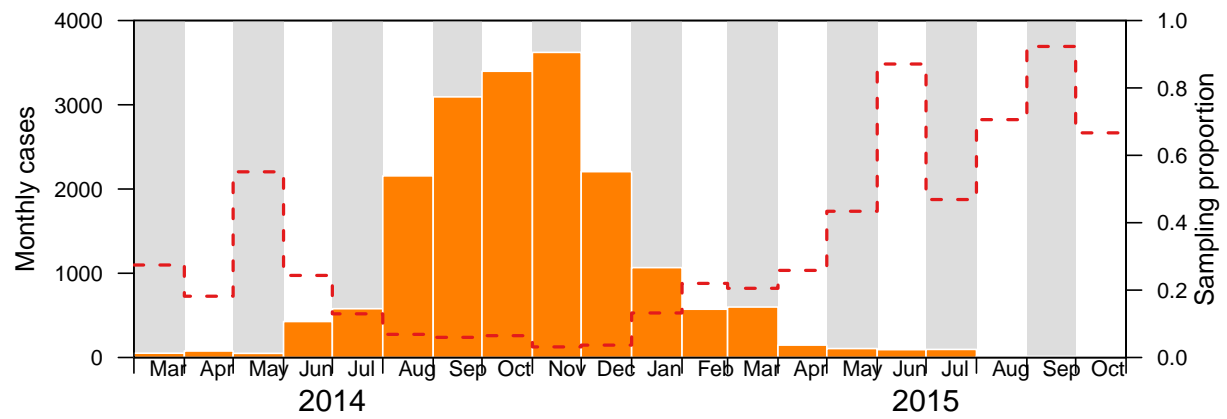
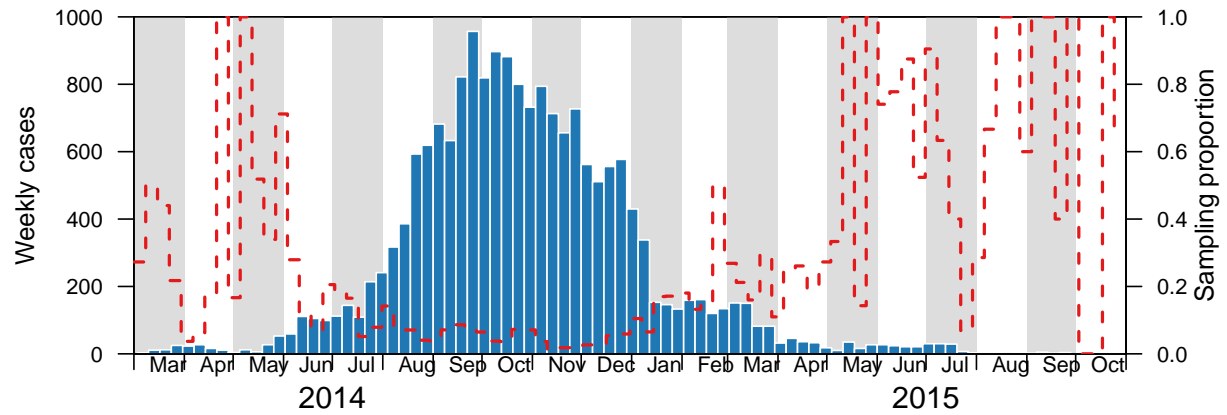
1	Summary	1
2	Datasets	2
2.1	Complete dataset	2
2.2	Southeast Sierra Leone	3
2.3	Western Sierra Leone	4
2.4	Eastern Sierra Leone	5
2.5	Liberia	6
2.6	Early outbreak	7
2.7	Subsampled complete outbreak	8
2.8	Subsampled complete outbreak big	9

1 Summary

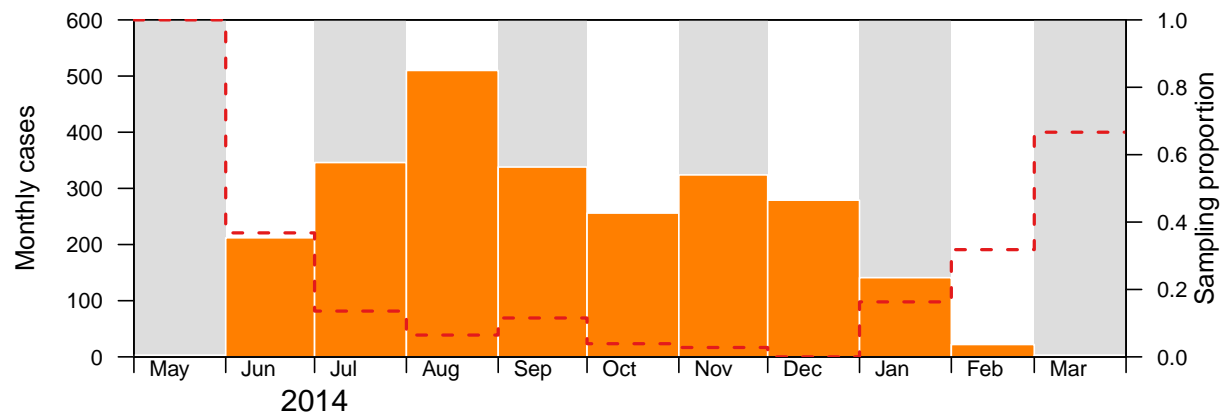
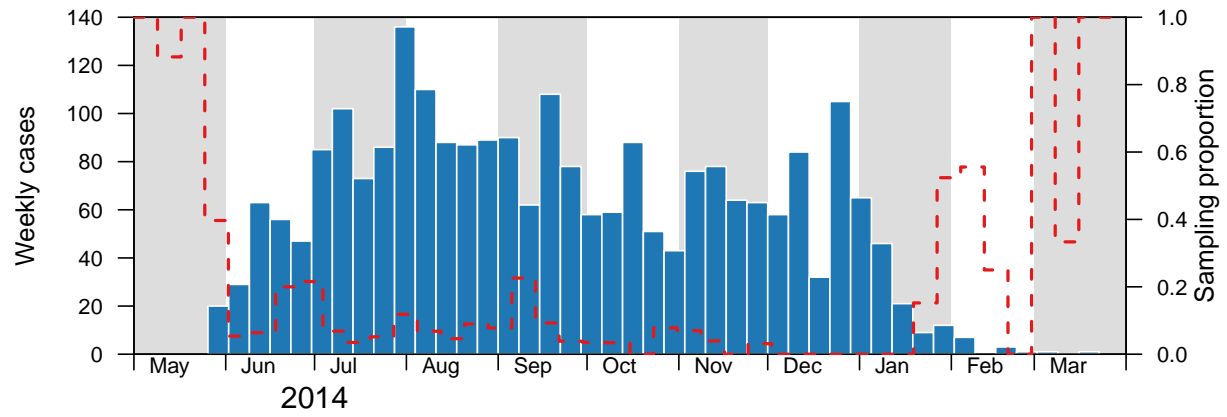
Extract the weekly and monthly empirical sampling proportions for the datasets extracted in `extract_datasets.rmd` using surveillance data from the epidemic. The empirical sampling proportion for a time period is defined as the number of sequences generated divided by the number of cases reported.

2 Datasets

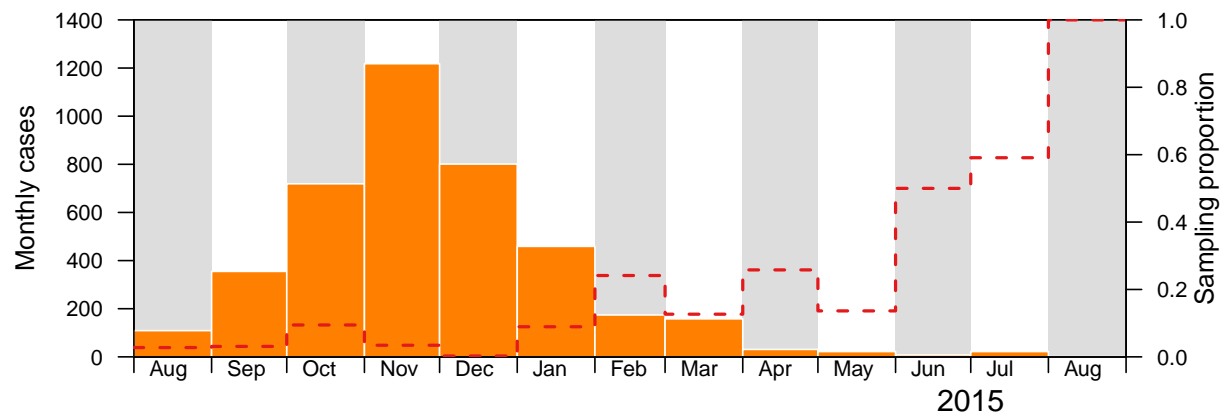
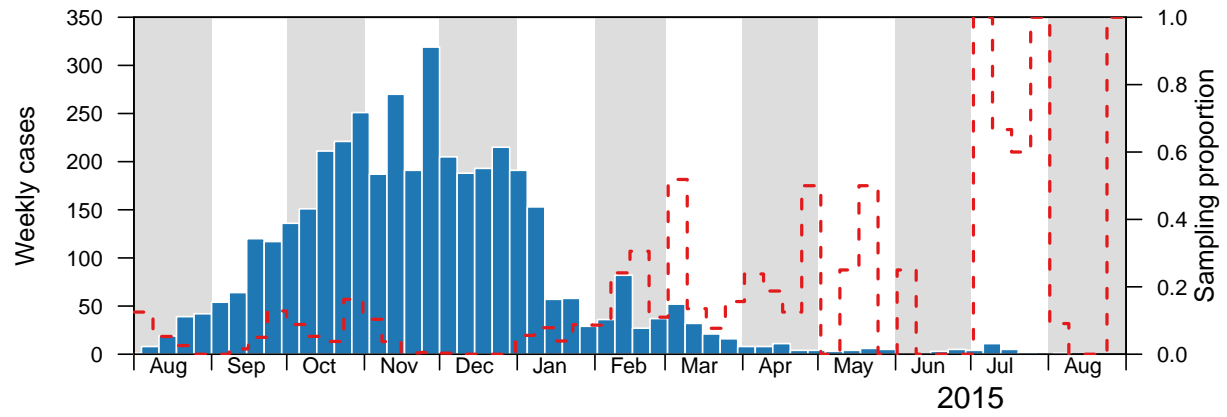
2.1 Complete dataset



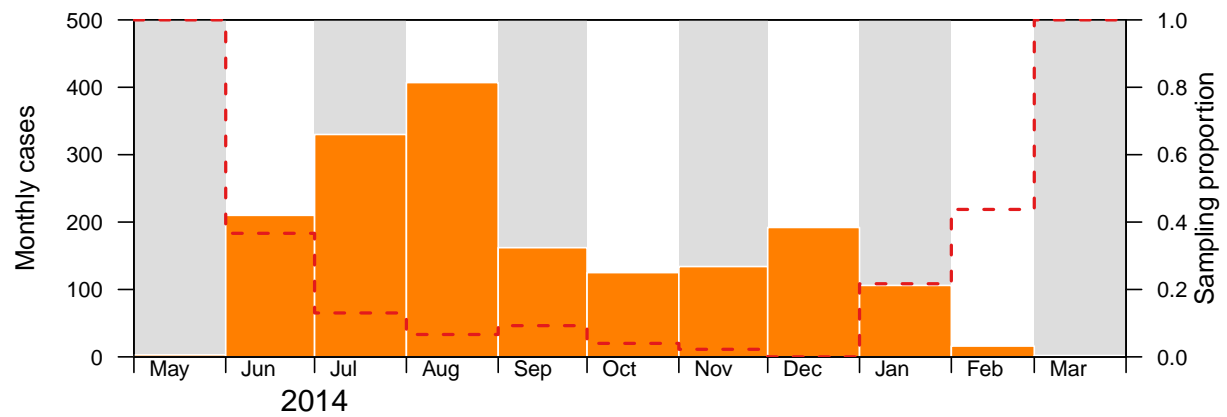
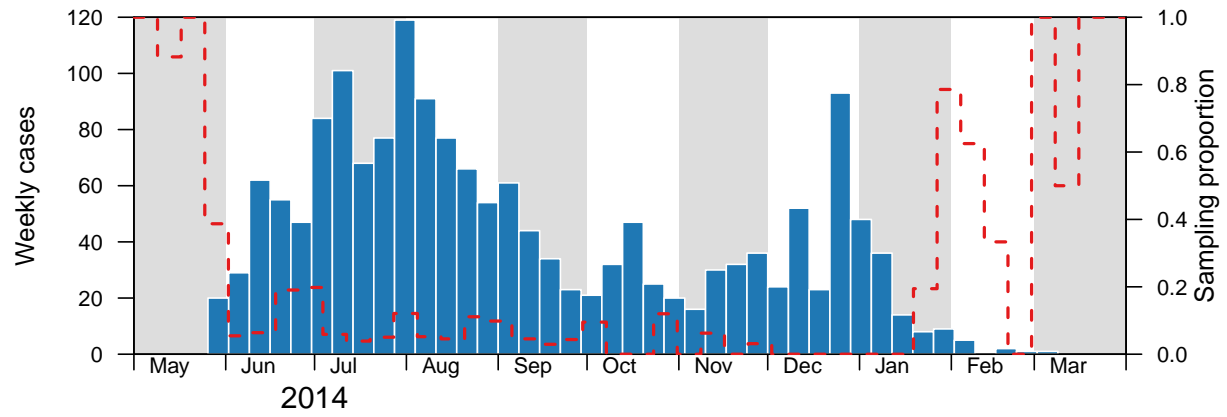
2.2 Southeast Sierra Leone



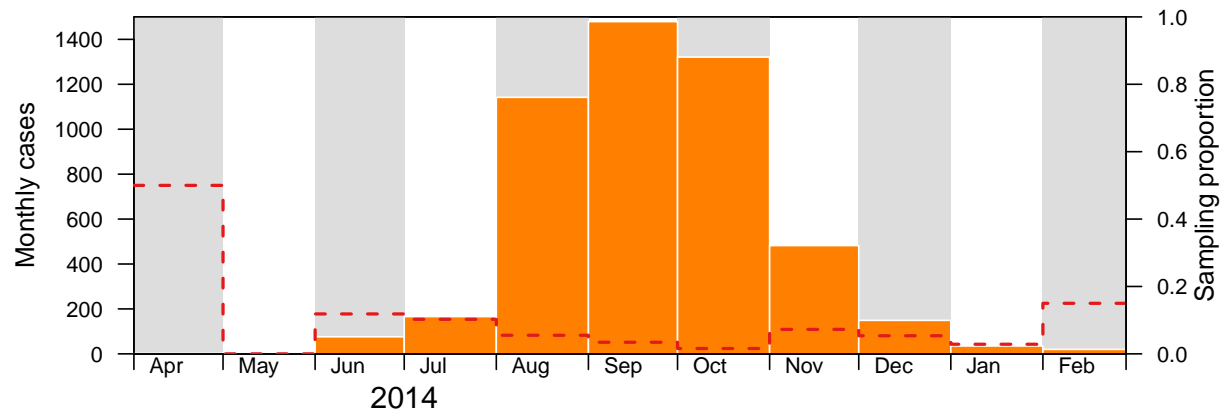
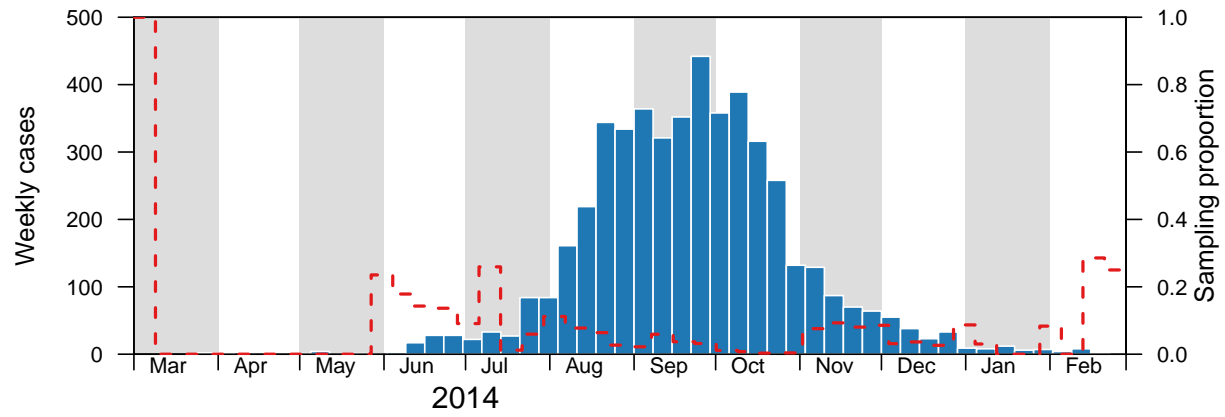
2.3 Western Sierra Leone



2.4 Eastern Sierra Leone

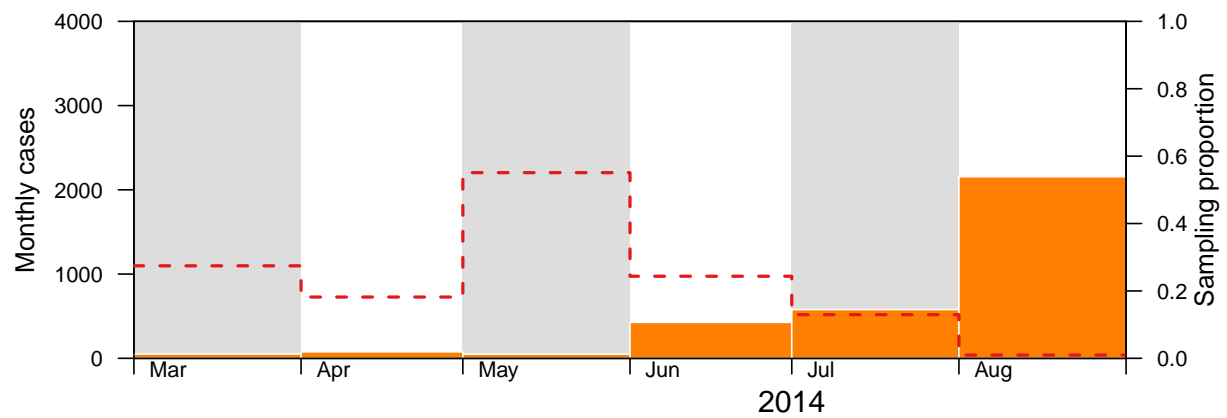
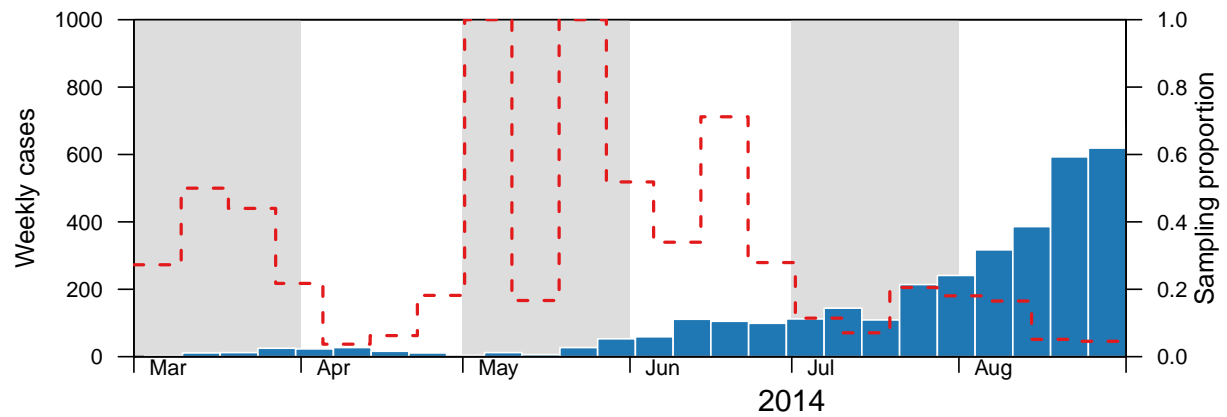


2.5 Liberia



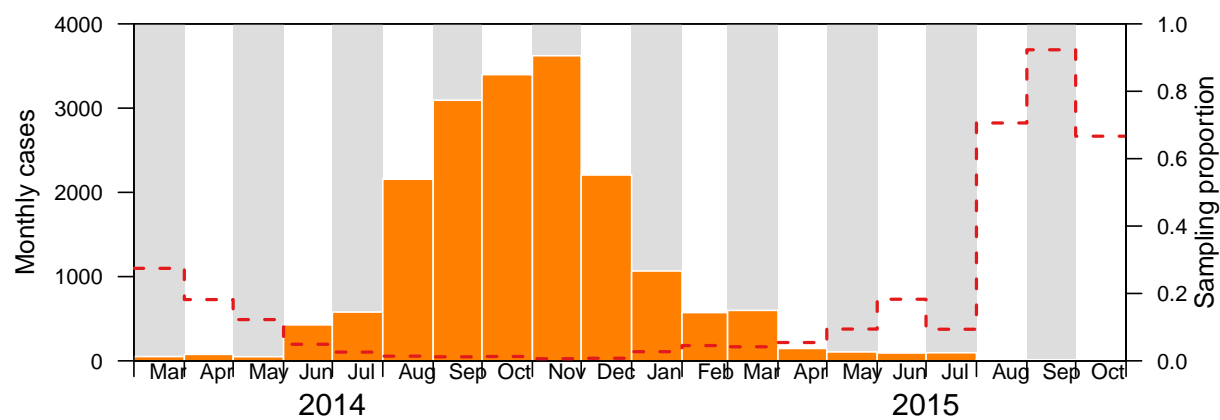
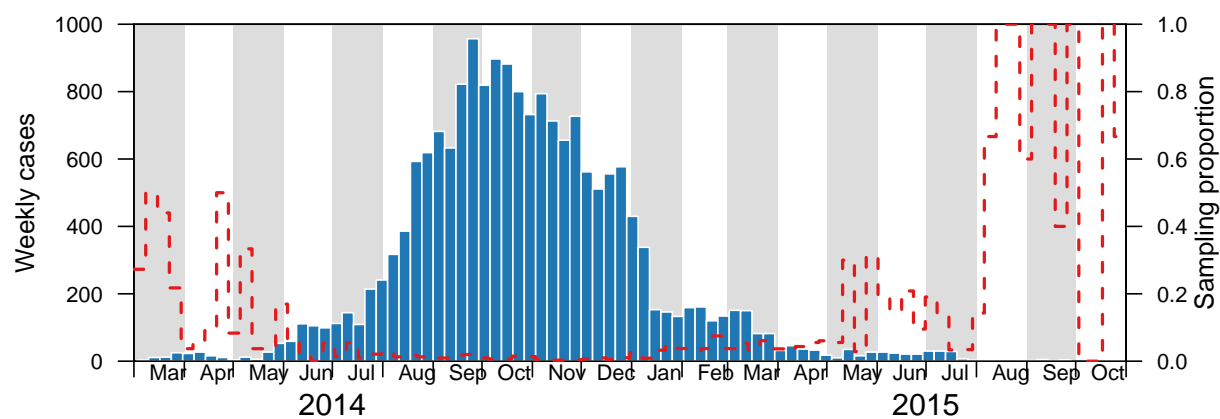
2.6 Early outbreak

All sequences until 4 August (WHO declares public health emergency).



2.7 Subsampled complete outbreak

Complete outbreak subsampled to approximately a quarter of sequences and roughly stratified by month.



2.8 Subsampled complete outbreak big

Complete outbreak subsampled to approximately half of all sequences and roughly stratified by month.

