

Exercises - SOLUTIONS

- The following is the best model I found.

```
wm<-gam(price~s(h.rain)+s(s.temp)+s(h.temp)+s(year),  
data=wine,family=Gamma(link=identity),gamma=1.4)  
plot(wm,pages=1,residuals=TRUE,pch=1,scale=0)  
acf(residuals(wm))  
gam.check(wm)  
predict(wm,wine,se=TRUE)
```

- A Gamma family seems to produce acceptable residual plots. `w.temp` appears to be redundant. Two way interactions between the weather variables only make the GCV and AIC scores worse. `gamma=1.4` is a prudent defense against overfitting, given that the sample size is so small. The effects are easy to interpret: (i) more recent vintages are worth less than older ones; (ii) low harvest rainfall and high summer temperatures are both associated with higher prices; (iii) there is some suggestion of an optimum harvest temperature at 17.5C, but the possible increases at very low and very high harvest temperatures make interpretation difficult (however, this harvest effect substantially increases the proportion deviance explained).