Tear-down approach

Emily Burchfield

Below is the tear-down approach for the CA dataset. All variables were initially included, and based on their significance they were slowly removed from the model to improve fit.

The null model was estimated as:

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: delta_tvp ~ 1 + (1 | HUC12)
##    Data: data
##
##      AIC      BIC   logLik deviance df.resid
## -25202.1 -25175.2  12604.0 -25208.1    56775
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.0219 -0.3672 -0.0123  0.3519  5.2901
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  HUC12    (Intercept) 0.00211  0.04593
##  Residual             0.03743  0.19347
## Number of obs: 56778, groups:  HUC12, 56
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.062427   0.006487  -9.624
```

And the ICC for the null model was found to be 0.0533592. The results from a GLM model show high significance due to the large sample size:

```
##
## Call:
## glm(formula = delta_tvp ~ delta_lc + diverse + Avg_WSEL_5yrChange +
##     WR_density + Perc_Rip + Perc_Pre1914 + GW_dnsty15, data = ds)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.19482  -0.07798  -0.00389   0.06986   1.03173
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -0.1309511  0.0040292 -32.501  < 2e-16 ***
## delta_lc            0.0072175  0.0017380   4.153 3.29e-05 ***
## diverse             0.0219983  0.0026135   8.417  < 2e-16 ***
## Avg_WSEL_5yrChange -0.0010032  0.0001678  -5.978 2.27e-09 ***
## WR_density          0.0010692  0.0004656   2.297   0.0216 *
## Perc_Rip            0.0207509  0.0070834   2.929   0.0034 **
## Perc_Pre1914       -0.0097268  0.0074626  -1.303   0.1924
## GW_dnsty15          0.2348366  0.0236895   9.913  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for gaussian family taken to be 0.03882261)
##
##     Null deviance: 2220.1  on 56441  degrees of freedom
## Residual deviance: 2190.9  on 56434  degrees of freedom
##   (350 observations deleted due to missingness)
## AIC: -23181
##
## Number of Fisher Scoring iterations: 2
```

Our first multi-level model includes all of the predictor variables as well as interaction terms between the i-level variable, a land-use flag, and group-level factors. The idea here is that land use changes could interact with HUC-level dynamics, such as the density of water rights and groundwater use. We also allow the effects of the i-level variable, `delta_lc`, to vary across groups.

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## delta_tvp ~ delta_lc + diverse + Avg_WSEL_5yrChange + WR_density +
##     Rip + P1914 + GW_dnsty15 + delta_lc * diverse + delta_lc *
##     WR_density + delta_lc * GW_dnsty15 + (1 + delta_lc | HUC12)
##    Data: ds
##
##      AIC      BIC   logLik deviance df.resid
## -25443.9 -25309.7  12736.9 -25473.9    56763
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.1003 -0.3654 -0.0106  0.3520  5.2784
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  HUC12    (Intercept) 0.001934 0.04398
##           delta_lc    0.001002 0.03166  -0.25
##  Residual             0.037200 0.19287
## Number of obs: 56778, groups:  HUC12, 56
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)         -0.1129422  0.0238844  -4.729
## delta_lc             0.0501641  0.0179112   2.801
## diverse              0.0176914  0.0118897   1.488
## Avg_WSEL_5yrChange  -0.0006469  0.0012186  -0.531
## WR_density           0.0104411  0.0205479   0.508
## Rip                 -0.0038122  0.0205282  -0.186
## P1914               -0.0099268  0.0088054  -1.127
## GW_dnsty15           0.3382996  0.1332400   2.539
## delta_lc:diverse    -0.0230103  0.0094008  -2.448
## delta_lc:WR_density  0.0013935  0.0011008   1.266
## delta_lc:GW_dnsty15 -0.0939994  0.1008586  -0.932
##
## Correlation of Fixed Effects:
##             (Intr) dlt_lc divers A_WSEL WR_dns Rip    P1914  GW_d15 dlt_l:
## delta_lc    -0.207
## diverse     -0.921  0.213
## Avg_WSEL_5C  0.392  0.012 -0.266
## WR_density  -0.401 -0.017  0.301 -0.106
## Rip          0.342  0.011 -0.264  0.077 -0.947
```

```
## P1914        0.350  0.024 -0.256  0.119 -0.673  0.403
## GW_dnsty15   0.275 -0.047 -0.480  0.191 -0.134  0.132  0.086
## dlt_lc:dvrs  0.198 -0.929 -0.246 -0.011  0.020 -0.014 -0.020  0.120
## dlt_lc:WR_d  0.004 -0.046  0.031  0.003 -0.021  0.000 -0.051 -0.038 -0.031
## dlt_l:GW_15 -0.048  0.184  0.128 -0.003 -0.012  0.010  0.004 -0.296 -0.425
##             d_:WR_
## delta_lc
## diverse
## Avg_WSEL_5C
## WR_density
## Rip
## P1914
## GW_dnsty15
## dlt_lc:dvrs
## dlt_lc:WR_d
## dlt_l:GW_15  0.081
```

We calculate the deviance for each model, which is a measure of model fit. We compare the deviance of the more complex model (`M1`) to less complex models progressively to see if dropping parameters improves fit. Since we are typically changing only one degree of freedom, we are looking fofr changes in deviance that are above ~3.8.

```
devcomp = getME(M1,"devcomp")
devM1 = as.numeric(devcomp$cmp[8])
```

The deviance for the full model is `devM1`. We played with dropping different interaction effects and found that only `delta_lc*diverse` significantly improves model fit. For reasons of parsimony, we dropped the other two interactions effects, `delta_lc*WR_density` and `delta_lc*GW_dnsty15`.

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## delta_tvp ~ delta_lc + diverse + Avg_WSEL_5yrChange + WR_density +
##     P1914 + Rip + GW_dnsty15 + delta_lc * diverse + (1 + delta_lc |
##     HUC12)
##    Data: ds
##
##      AIC      BIC   logLik deviance df.resid
## -25445.2 -25328.9  12735.6 -25471.2    56765
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.1002 -0.3654 -0.0103  0.3525  5.2796
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  HUC12    (Intercept) 0.001925 0.04388
##           delta_lc    0.001055 0.03248  -0.24
##  Residual             0.037201 0.19287
## Number of obs: 56778, groups:  HUC12, 56
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)        -0.1143902  0.0238154  -4.803
## delta_lc            0.0549469  0.0179248   3.065
## diverse             0.0188254  0.0117685   1.600
## Avg_WSEL_5yrChange -0.0006542  0.0012164  -0.538
## WR_density          0.0108760  0.0205204   0.530
```

```
## P1914              -0.0093215  0.0087866  -1.061
## Rip                -0.0037115  0.0205062  -0.181
## GW_dnsty15          0.3050330  0.1270602   2.401
## delta_lc:diverse   -0.0269094  0.0086834  -3.099
##
## Correlation of Fixed Effects:
##             (Intr) dlt_lc divers A_WSEL WR_dns P1914  Rip    GW_d15
## delta_lc    -0.202
## diverse     -0.924  0.196
## Avg_WSEL_5C  0.393  0.012 -0.269
## WR_density  -0.402 -0.016  0.306 -0.106
## P1914        0.352  0.021 -0.258  0.119 -0.675
## Rip          0.343  0.010 -0.267  0.077 -0.947  0.403
## GW_dnsty15   0.274  0.007 -0.467  0.199 -0.144  0.091  0.141
## dlt_lc:dvrs  0.196 -0.957 -0.213 -0.013  0.017 -0.021 -0.011 -0.006
```

This suggests that diversity moderates the effect of land use change in a pixel, where increased diversity reduces the effect of landuse change on TVP.

The change in deviance is `devstat`. In `M3` we drop `Rip` due to low significance.

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## delta_tvp ~ delta_lc + diverse + Avg_WSEL_5yrChange + WR_density +
##     P1914 + GW_dnsty15 + delta_lc * diverse + (1 + delta_lc |      HUC12)
##    Data: ds
##
##      AIC      BIC   logLik deviance df.resid
## -25447.2 -25339.8  12735.6 -25471.2    56766
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.1002 -0.3654 -0.0103  0.3525  5.2797
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  HUC12    (Intercept) 0.001930 0.04393
##           delta_lc    0.001054 0.03247  -0.24
##  Residual             0.037201 0.19287
## Number of obs: 56778, groups:  HUC12, 56
##
## Fixed effects:
##                     Estimate Std. Error t value
## (Intercept)        -0.112896   0.022390  -5.042
## delta_lc            0.054975   0.017918   3.068
## diverse             0.018251   0.011351   1.608
## Avg_WSEL_5yrChange -0.000637   0.001214  -0.525
## WR_density          0.007357   0.006602   1.114
## P1914              -0.008678   0.008044  -1.079
## GW_dnsty15          0.308246   0.125923   2.448
## delta_lc:diverse   -0.026925   0.008680  -3.102
##
## Correlation of Fixed Effects:
##             (Intr) dlt_lc divers A_WSEL WR_dns P1914  GW_d15
## delta_lc    -0.218
## diverse     -0.919  0.206
## Avg_WSEL_5C  0.391  0.012 -0.258
```

```
## WR_density  -0.255 -0.022  0.171 -0.103
## P1914        0.248  0.019 -0.170  0.097 -0.996
## GW_dnsty15   0.242  0.006 -0.450  0.191 -0.032  0.037
## dlt_lc:dvrs  0.213 -0.957 -0.224 -0.012  0.020 -0.018 -0.005
```

```
## [1] 0.03262752
```

This doesn't really make a difference, deviance change is `devstat`, but we'll drop `Rip` since we include a flag for `P1914`. Next we drop `Avg_WSEL_5yrChange` which isn't significant and doesn't measure water use in terms of density. To be consistent, we drop this in the next model.

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## delta_tvp ~ delta_lc + diverse + WR_density + P1914 + GW_dnsty15 +
##     delta_lc * diverse + (1 + delta_lc | HUC12)
##    Data: ds
##
##      AIC      BIC   logLik deviance df.resid
## -25448.9 -25350.5  12735.5 -25470.9    56767
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.1001 -0.3655 -0.0102  0.3526  5.2797
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  HUC12    (Intercept) 0.001941 0.04406
##           delta_lc    0.001054 0.03246  -0.24
##  Residual             0.037201 0.19287
## Number of obs: 56778, groups:  HUC12, 56
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)      -0.108246   0.020663  -5.239
## delta_lc          0.055163   0.017917   3.079
## diverse           0.016675   0.011000   1.516
## WR_density        0.006998   0.006583   1.063
## P1914            -0.008267   0.008026  -1.030
## GW_dnsty15        0.321476   0.124165   2.589
## delta_lc:diverse -0.027016   0.008679  -3.113
##
## Correlation of Fixed Effects:
##             (Intr) dlt_lc divers WR_dns P1914  GW_d15
## delta_lc    -0.237
## diverse     -0.920  0.212
## WR_density  -0.235 -0.021  0.150
## P1914        0.230  0.018 -0.151 -0.996
## GW_dnsty15   0.186  0.003 -0.423 -0.012  0.019
## dlt_lc:dvrs  0.232 -0.957 -0.231  0.020 -0.017 -0.002
```

```
## [1] 0.2722474
```

No significant change in deviance with this change, deviance is `devstat`, but for parsimony and consistency, we drop this variable. In the final model, we look at the effect of removing the random effect on `delta_lc`.

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula:
## delta_tvp ~ delta_lc + diverse + WR_density + P1914 + GW_dnsty15 +
##     delta_lc * diverse + (1 | HUC12)
##    Data: ds
##
##      AIC      BIC   logLik deviance df.resid
## -25231.6 -25151.1  12624.8 -25249.6    56769
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.0143 -0.3707 -0.0105  0.3512  5.2831
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  HUC12    (Intercept) 0.001712 0.04137
##  Residual             0.037409 0.19342
## Number of obs: 56778, groups:  HUC12, 56
##
## Fixed effects:
##                   Estimate Std. Error t value
## (Intercept)      -0.103933   0.019481  -5.335
## delta_lc          0.040062   0.007128   5.621
## diverse           0.013915   0.010385   1.340
## WR_density        0.006990   0.006471   1.080
## P1914            -0.008194   0.007890  -1.039
## GW_dnsty15        0.294704   0.119677   2.463
## delta_lc:diverse -0.018064   0.003298  -5.478
##
## Correlation of Fixed Effects:
##             (Intr) dlt_lc divers WR_dns P1914  GW_d15
## delta_lc    -0.097
## diverse     -0.919  0.085
## WR_density  -0.244 -0.011  0.156
## P1914        0.238  0.009 -0.157 -0.996
## GW_dnsty15   0.191  0.012 -0.432 -0.014  0.021
## dlt_lc:dvrs  0.100 -0.970 -0.098  0.010 -0.009 -0.012

## [1] 221.35
```

There's a huge change in deviance when we drop the random effect, suggesting we should keep it in the model. The final model `M4` shows significant effects for groundwater, the interaction between `delta_lc*diverse` and slight effects of `diverse`. We played with dropping the two water rights variables one at a time, and in all cases, significance was very small. We could drop water rights from the analysis, but theoretically it's interesting to include, to show that the structure of surface water water rights matters less than groundwater use. In any case, leaving them in doesn't reduce the overall fit of the model to the data.