New question: What should I do when I've measured the same e.u. multiple times?

Examples:

- physical therapy: measure of joint flexibility at 0, 1, 2, 3, $\cdots$ 10 weeks of therapy
- bioenergy study: 3 substrates, RCBD,
  measure amount of product produced after 1, 2, 3, 4 days.
- animal nutrition: 3 diets, RCBD, measure weight at week 0, 1, 2, 4, 6, 8
- corn root worm: 2 insecticide treatments, damage measured 10 times
- asparagus: perennial crop, 4 harvesting intensities, RCBD,
  yield measured in years 2, 3, and 4.

An important distinction, are different eu's measured each time, or are they the same eu?

- Different eu's: Each eu is measured once. Presumably sampling time randomly assigned.
  - Corn root worm: Measurements are destructive. A plant can't be measured twice. damage over time. 2 way factorial treatment structure:
  - No new issues: 2 way factorial, insecticide x meas. time. independent errors
- Same eu's: More common, more efficient use of resources, since eu's are often expensive
  - New issues: observations over time are probably correlated.
  - Repeated measures analysis: accounts for that correlation
  - Asparagus cutting study: plot randomly assigned to a treatment. Same trt for all years
    - Each plot repeated measured
    - A plot with a higher than average yield in year 2 is likely to be above average in years 3 and 4.
    - Analysis needs to account for the potential correlation within an e.u.
  - Most common: treatments randomly assigned once, at start of study.
  - No new randomization of treatments each year
    - Reusing a field for new experiments each year does not create repeated measures data
    - So long as not comparing between years

Different sorts of studies with same eu measured multiple times

- Longitudinal study:

  - Subjects randomized to treatment at start
  - Measured multiple times. May be regular or irregular intervals.

- Crossover study, usually only two treatments, perhaps 3.
  - Subjects randomized to initial treatment
  - Half way through, switch to 2nd treatment.
  - Commonly, with a "washout" period in between.
  - So half the subjects get A followed by B; other half get B followed by A.
  - May be measured only at end of each treatment period (one obs per subject and treatment)
  - or multiple times throughout each period.

- Two sizes of eu: longitudinal + crossover
  - Subjects randomly assigned to one level of factor A and to order of levels of factor B
  - A is constant over time (as in a longitudinal study)
  - B changes over time (as in a crossover study)

Focus (for now) on longitudinal studies. Analysis options:

- Compute a summary statistic for each eu.
  - One number for each eu.
  - Some possible choices:
    * Value at last observation time, best when impact is cumulative
    * Average all observations, best when impact is sharp or total is relevant
    * Compute slope over time for each subject
  - Should be done more often. Simplifies the analysis. See Murdoch paper

- A variation on this:
  - Separately analyze each year's data
  - Summary statistic = yield at year 1 for analysis 1
  - Yield at year 2 for analysis 2, yield at year 3 for analysis 3
  - Simple, effective, well known, but no way to compare between years.

- "Split-plot in time"
  - Recognize two sizes of "eu":

  - subject assigned to treatment
  - time "assigned" to observation
  - implicitly assumes same correlation between all times
    * Nutrition study: same correlation between day 1 and 3 as between day 1 and 21
    * If correlation decays with time, seems unreasonable

- Model correlation as function of time lag
  - Correlation between day 1 and 3 larger than that between day 1 and 21
  - But correlation between day 1 and 3 same as between day 3 and 5
  - Only a function of difference between two times (the lag).
  - Various possible models

- Model correlation as a function of starting and stopping time
  - Correlation between day 1 and 3 larger than that between day 1 and 21
  - And correlation between day 1 and 3 may differ from that between day 3 and 5
  - And, as usually implemented, variance on day 1 may differ from that on day 3 or day 5.
  - My (and others's) experience is this 'allows" too much.
  - Most data sets have more consistencies than this. E.g. constant variance

- Treat all responses from an eu as a vector valued response
  - Analyze using MANOVA: Multivariate ANOVA
  - Allows arbitrary variance-covariance matrix
    * correlations can very between intervals
    * variances at each time may not be the same
  - Each treatment has a mean vector
  - So null hypothesis is very general: no difference in year 2 and no difference in year 3, ...
  - And MANOVA rejects Ho if there is a difference **anywhere**
  - but doesn't tell you which times (responses) are different
  - So you end up doing separate analyses for each time to find that out.
  - I don't recommend using MANOVA:
    * The null hypothesis is dumb - too general
    * The analysis is very sensitive to assumptions, e.g. same VC matrix for all treatments
    * Tends to overfit the VC matrix (lots of parameters there if many times)

    * Missing one observation (time) on a subject leads to dropping all data for that subject

Modeling the correlation between observations make on the same e.u.

Basic idea: find a simple but reasonable model for correlations between observations made at different times on the same eu. Lots of possible models:

- ar(1): autoregressive order 1. 1 parameter for correlation, ρ, 1 for variance.
  - correlation declines to 0 with time lag between to observations
  - correlation between 1 and 2, or 3 and 4: ρ
  - correlation between 1 and 3, or 2 and 4: ρ^2
  - correlation between 1 and 4: ρ^3
- ar(1) + random effect: like ar(1), except that correlation declines to a non-zero value
  - Not very popular, but I often find this is fits better than any other model.
- arh(1): ar(1) with different variances for each time point
- and its arh(1) + re extension
- ante-dependence: allows correlation between 1 and 2 to differ from 2 to 3.
- unstructured: no pattern - every correlation and variance is different

Here are all the details: covariance matrices for three repeated measures per e.u.
Remember, a covariance matrix includes the variance of each observation.
The associated correlation matrix does not.

Variance-covariance matrices:

Independence
1 parameter: σ^2

$$\begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$$

Split-plot, Compound Symmetry
2 parameters: σ_m^2, σ_s^2

$$\begin{bmatrix} \sigma_m^2+\sigma_s^2 & \sigma_m^2 & \sigma_m^2 \\ \sigma_m^2 & \sigma_m^2+\sigma_s^2 & \sigma_m^2 \\ \sigma_m^2 & \sigma_m^2 & \sigma_m^2+\sigma_s^2 \end{bmatrix}$$

AR(1)
2 parameters: σ^2, ρ

$$\begin{bmatrix} \sigma^2 & \rho\,\sigma^2 & \rho^2\,\sigma^2 \\ \rho\,\sigma^2 & \sigma^2 & \rho\,\sigma^2 \\ \rho^2\,\sigma^2 & \rho\,\sigma^2 & \sigma^2 \end{bmatrix}$$

AR(1) + RE
3 parameters: σ_m^2, σ^2, ρ

$$\begin{bmatrix} \sigma_m^2+\sigma^2 & \sigma_m^2+\rho\,\sigma^2 & \sigma_m^2+\rho^2\,\sigma^2 \\ \sigma_m^2+\rho\,\sigma^2 & \sigma_m^2+\sigma^2 & \sigma_m^2+\rho\,\sigma^2 \\ \sigma_m^2+\rho^2\,\sigma^2 & \sigma_m^2+\rho\,\sigma^2 & \sigma_m^2+\sigma^2 \end{bmatrix}$$

ARH(1)
4 parameters: σ_1^2, σ_2^2, σ_3^2, ρ

$$\begin{bmatrix} \sigma_1^2 & \rho\,\sigma_1\sigma_2 & \rho^2\,\sigma_1\sigma_3 \\ \rho\,\sigma_1\sigma_2 & \sigma_2^2 & \rho\,\sigma_2\sigma_3 \\ \rho^2\,\sigma_1\sigma_3 & \rho\,\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}$$

ARH(1) + RE
5 parameters: σ_m^2, σ_1^2, σ_2^2, σ_3^2, ρ

$$\begin{bmatrix} \sigma_1^2+\sigma_m^2 & \rho\,\sigma_1\sigma_2+\sigma_m^2 & \rho^2\,\sigma_1\sigma_3+\sigma_m^2 \\ \rho\,\sigma_1\sigma_2+\sigma_m^2 & \sigma_2^2+\sigma_m^2 & \rho\,\sigma_2\sigma_3+\sigma_m^2 \\ \rho^2\,\sigma_1\sigma_3+\sigma_m^2 & \rho\,\sigma_2\sigma_3+\sigma_m^2 & \sigma_3^2+\sigma_m^2 \end{bmatrix}$$

ANTE(1)
5 parameters: σ_1^2, σ_2^2, σ_3^2, ρ_1, ρ_2

$$\begin{bmatrix} \sigma_1^2 & \rho_1\,\sigma_1\sigma_2 & \rho_1\rho_2\,\sigma_1\sigma_3 \\ \rho_1\,\sigma_1\sigma_2 & \sigma_2^2 & \rho_2\,\sigma_2\sigma_3 \\ \rho_1\rho_2\,\sigma_1\sigma_3 & \rho_2\,\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}$$

UN
6 parameters: σ_1^2, σ_2^2, σ_3^2, σ_{12}, σ_{23}, σ_{13}

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}$$

Corresponding correlation models

Independence

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Split-plot

$$\begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$$

AR(1)

$$\begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$$

AR(1) + RE

$$\begin{bmatrix} 1 & f+(1-f)\,\rho & f+(1-f)\,\rho^2 \\ f+(1-f)\,\rho & 1 & f+(1-f)\,\rho \\ f+(1-f)\,\rho^2 & f+(1-f)\,\rho & 1 \end{bmatrix}$$

ARH(1)

$$\begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$$

ARH(1) + RE

not nice

ANTE(1)

$$\begin{bmatrix} 1 & \rho_1 & \rho_1\rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_1\rho_2 & \rho_2 & 1 \end{bmatrix}$$

UN

$$\begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}$$

Summary of characteristics of correlation models

Model	Variances	Correlations
Independence	constant	zero, for all pairs of observations
Split-plot	constant	same for all pairs
AR(1)	constant	decline with separation
		eventually zero
		same for 1 to 2 and 2 to 3
AR(1) + RE	constant	decline to non-zero
ARH(1)	different	decline with separation
ARH(1) + RE	different	decline to non-zero
ANTE(1)	different	decline to zero
		different for 1 to 2 and 2 to 3
UN	different	every pair different

How to choose an appropriate correlation structure?

I suggest three possible ways.

- Your understanding of the data: what is reasonable?
- What has been used in other studies?
- What do these data suggest is a reasonable model?
  - Plot correlations between pairs of observations - look at the pattern(s)
  - Fit a model and assess how well it fits
  - AIC and BIC statistics
    - used just like you used for variable selection in multiple regression
    - AIC $= -2 \log \text{likelihood} + 2p$
    - likelihood: measures how well data fit a model. Generalizes sum-of-squares.
    - concept: fit + penalty for complexity
    - p: number of parameters in the model
    - BIC $= -2 \log \text{likelihood} + p \log N$
    - $\log N > 2$, for all but the smallest data sets, so larger penalty
  - Find the model with the smallest AIC or BIC value
  - If REML, need to use the same fixed effect model
  - REML should only be used to evaluate the variance/covariance structure
  - AICc includes a small sample size "correction"
    prefered to AIC, but usually not any different,
    unless very small numbers of observations
    In which case, any model selection criterion is suspect.
  - BIC has a higher penalty for complexity: if different from AIC/AICc, BIC will choose simpler model

For the asparagus data:

Model	# param	AIC	AICc	BIC
Independence	1	342.0	342.1	343.5
Split-plot	2	337.0	337.4	338.5
AR(1)	2	**333.1**	**333.5**	**334.6**
AR(1)+RE	3	$\hat{\sigma}^2_m = 0$, so same as AR(1)		
ARH(1)	4	335.8	337.2	338.9
ARH(1)+RE	5	$\hat{\sigma}^2_m = 0$, so same as ARH(1)		
ANTE(1)	5	335.7	338.9	340.3
UN	6	335.7	338.9	340.3

ar(1) is the "best" model among those evaluated here.

Effect of correlation structure on results for asparagus study

Model	Trt A - Trt B		Year 2 - Year 3	
	s.e.	d.f.	s.e.	d.f.
Independence	12.09	33	10.46	33
Split-Plot	17.84	9	7.82	24
AR(1)	18.34	9.3	6.98	23.1
ARH(1)	18.01	8.53	6.21	20.6
ANTE(1)	18.48	8.74	7.64	12
UN	18.62	8.29	7.49	12

Assuming independence over time produces clearly different results. **clearly wrong!** Too small a se for averages of over times, e.g. differences between treatments. Too large a se for comparisons between times.

Choice of model for correlated data has much less impact on results. Models with lots of parameters, e.g. UN, ANTE(1), have smaller d.f. and larger s.e. This is a general pattern. Results in loss of power. Multivariate approach corresponds to UN; also lower power.

Two rep. meas. models with similar AIC and BIC values will usually (but not always) produce similar (but not same) conclusions about treatment effects (F statistics, p-values, s.e. d.f.).

My advice:

- Can you use a summary statistic to answer your questions? If so, do that and don't worry about choosing a correlation model
- If not, use AIC and biological knowledge to choose correlation model. Try to get close to a reasonable model. Don't worry too much about getting the model exactly right.

Final thoughts:

- When only 2 repeated measures, there is only one correlation coefficient. Choice of models irrelevant, because all will fit the same. The one difference is whether the two variances are the same or not.
- When 3 repeated measures, there are 3 correlation coefficients (1-2, 2-3, and 1-3). Not much room for models to differ. My experience is that simpler models are usually selected.
- When many repeated measures, there are many more correlation coefficients and much more room for models to differ.
- Remember that AIC/BIC are relative measures. They identify the best fitting among the models you have specified. The real best model may be something not in the set you are considering. In other words, AIC/ BIC will identify the best of a bad lot. That doesn't mean the selected model is the correct one.
- A residual vs predicted values plot may help you diagnose unequal variances. It will not help identify correlation patterns.
- It is always possible (code not provided) to compute the correlations between pairs of times and plot correlation against the time lag. That can help decide among potential models. And, help you decide whether the "best" model is still a bad fit.
- If in doubt about the correlation model, use all reasonable candidates and see if the things you really care about (e.g. treatment differences, time differences, or their interaction) change depending on the correlation model. My experience is that usually they do not, which is reassuring.