Analisis\_generico\_baza

# Cargar paquetes

# DATOS INSTAR

## Importar datos INSTAR

* La tabla de datos se construye en excel y se guarda como .csv
* Se abre con procesador de texto, y se comprueba que la fecha esta en formato mm-dd-yy
* Se guarda como .txt Las columnas que constituyen la tabla son: fecha, huevo, L1, L2, crisalida, radiacion, tmax, tmin, tmed y exergia.

## Anadir biociclos al data.frame

Aparece dos chunks para este proceso, uno para Baza y otro para Sierra Nevada, ya que tienen fechas de inicio diferentes.

## Agregacion INSTAR

Media del vigor y sumas (l1, l2 y l1+l2) por biociclo

# DATOS COPLAS

## Importar datos COPLAS

COPLAS<-read.csv("zona\_baza.csv", header=TRUE, sep=",")  
names(COPLAS)[1] <- "RODAL"

Activar el rodal necesario para el analisis

## Definir biociclo COPLAS

## Union de tablas INSTAR y COPLAS

En esta tabla aparecen los datos de Instar agregados, los Ln de sumL1, sumL2 y sumL1+sumL2, el anio y el grado de infestacion de COPLAS.

# VALIDACION EXTERNA

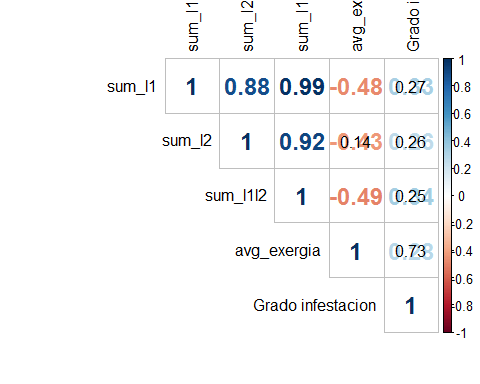
## Correlacion [L1, L2, L1+L2 vs. COPLAS]

Ya que es dificil explicar que una variable es dependiente de otra. La mejor forma de validar el modelo (con datos reales) es haciendo una correlacion. Con este analisis sabremos si existe relacion o no entre la variable Instar y COPLAS. Sabremos tambien la intensidad de esta relacion y si es significativa. Como nuestros datos no siguen una distribucion normal, haremos una correlacion con datos no parametricos: 'Spearman'

Funcion en R (pedir a Antonio este codigo) que plotea la relacion entre ellos con un circulo (cuanto mas relacionados, mas grande el circulo) y si es negativa o positiva. Da tambien el p valor de la correlacion. *"Puedo tener fuerte relacion pero el coef no es diferente de 0, a lo mejor con mas datos si."*"

Deberiamos esperar que haya mas relacion entre coplas y exergia, luego coplas y larvas. Cuando hay relacion negativa es que es inversa. Y como no sabemos quien es x e y, quiere decir que cuando una aumenta, otra disminuye, peor no sabemos quien aumente y quien disminuye.

#Plot  
corrplot(correlacion, p.mat = res1[[1]], sig.level=0.05, method = "number", type = "upper", tl.col = "black", number.cex = 1.5, insig = "p-value")



* La correlacion de exergia y coplas resulta ser menos fiable de los esperado. Posiblemente porque su modelizacion es simple.
* La correlacion frente a COPLAS de L1+L2 y L1 son parecidas. Esto quiere decir que L2 no aporta mucho (o nada) cuando se agregan. Esto probablemente sea porque la agregacion de las larvas indica la carga defoliadora en el sistema y al hacer esta agregacion (pobre) se le da la misma capacidad defoliadora a L1 que a L2, cuando en realidad L2 tiene mas.

# VALIDACION INTERNA

## ACP

Autocorrelation plot. Nos ayudara a saber si existen ciclos dentro de toda la serie temporal. ESTE ANALISIS SOLO SE HACE CON LOS DATOS DE BAZA

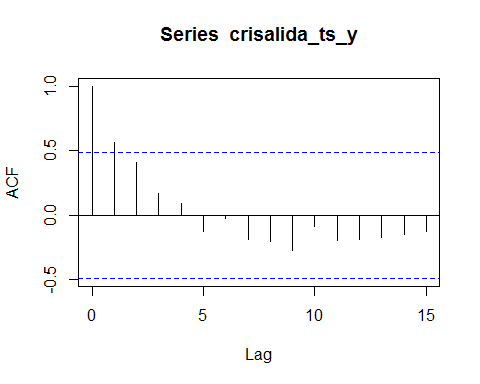
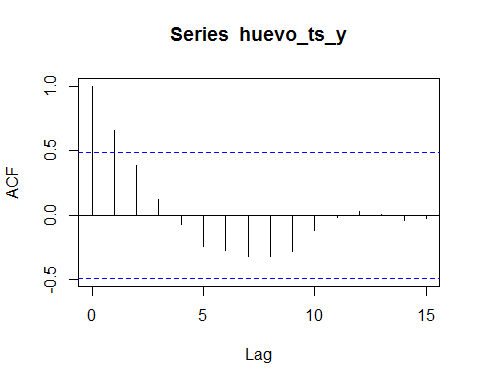
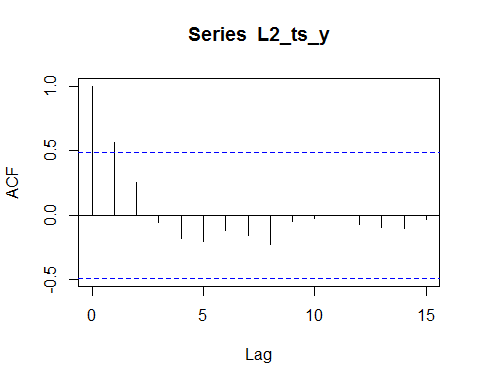
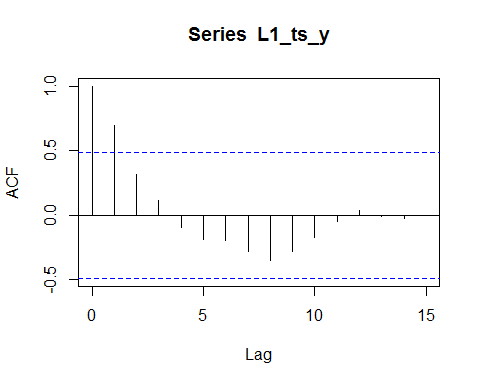
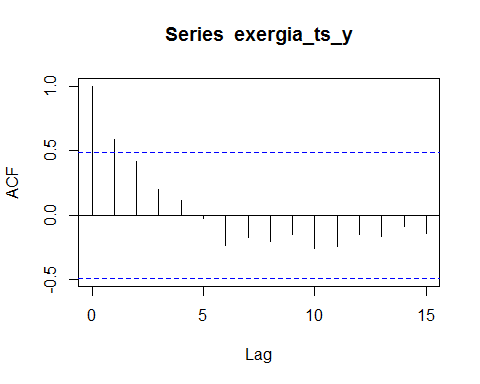
"*Those plots are showing you the correlation of the series with itself, lagged by x time unitscorrelation of the series with itself, lagged by x time units. So imagine taking your time series of length TT, copying it, and deleting the first observation of copy#1 and the last observation of copy#2. Now you have two series of length T for which you calculate a correlation coefficient. This is the value of of the vertical axis at x=1 in your plots. It represents the correlation of the series lagged by one time unit. You go on and do this for all possible time lags xx and this defines the plot*.

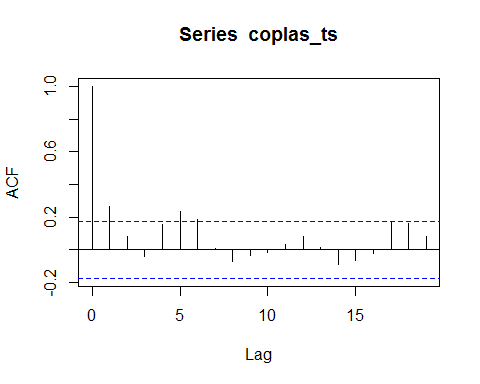
*The answer to your question of what is needed to report a pattern is dependent on what pattern you would like to report. But quantitatively speaking, you have exactly what I just described: the correlation coefficient at different lags of the series. You can extract these numerical values by issuing the command acf(x.ts,100)$acf*.

*In terms of what lag to use, this is again a matter of context. It is often the case that there will be specific lags of interest. Say, for example, you may believe the fish species migrates to and from an area every ~30 days. This may lead you to hypothesize a correlation in the time series at lags of 30. In this case, you would have support for your hypothesis*"

"*The blue lines give the values beyond which the autocorrelations are (statistically) significantly different from zero*"

Hay que definir el LAG. lag.max= Lo hemos definido como el numero de biociclos.

**CORRELOGRAMAS INSTAR** 

**CORRELOGRAMA COPLAS** 

Los Lags 1, 5, 6 y 17 son significativos. ?? La de 5 y 6 podria referirse al ciclo del que hablan en la literatura.

## Analisis de tendencia

ind\_feno<-read.csv("indicadores\_feno\_bz006.csv", header = TRUE, sep = ",", dec = ".")

library(Kendall)

Se eligen indicadores de cada variable para calcular su tendencia.

En las variables huevo, L1, L2 y crisalida:

* Fecha del maximo. *HECHO*
* Fecha del primer dia distinto a cero. *EN PROCESO*
* Fecha del ultimo dia distinto a cero. *EN PROCESO*

En la variable exergia:

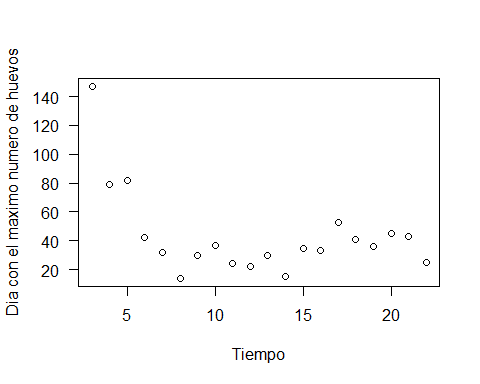
* Fecha del maximo.*HECHO*
* Fecha del minimo. *HECHO*

H0= no ha tendencia. Si p<0.05 entonces si hay tendencia!

**TENDENCIA EN HUEVOS**

*DIA DE PICO MAXIMO*

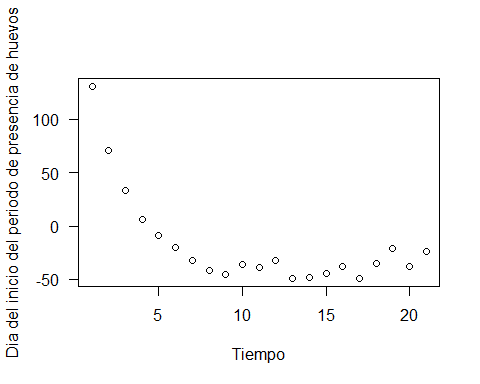
## Score = 33 , Var(Score) = 949  
## denominator = 189.4993  
## tau = 0.174, 2-sided pvalue =0.29891



Los valores del tau de MannKendall van de -1 a 1, por lo tanto un tau =0.174 no indicaria tendencia. Aunque no podriamos asegurarlo ya que el dato no es significativo. Muy marcada la diferencia de los 3 primeros datos con el resto, podra deberse al efecto llamado 'calentamiento dle modelo' Tal vez, podriamos decir que el modelo en el tiempo se va 'adecuando' a las fechas reales, si es que la fecha real del pico max de huevo es sobre el dia 40 (mediados de febrero).

*PRIMER DIA*

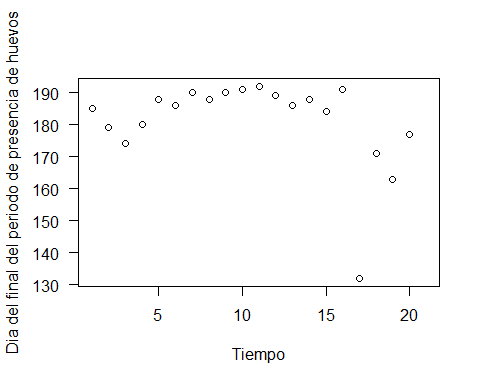
## Score = -89 , Var(Score) = 1093.667  
## denominator = 208.4946  
## tau = -0.427, 2-sided pvalue =0.0077916



Tendencia negativa (tau= -0.43) y significativa. Puede deberse al calentamiento dle modelo..

*ULTIMO DIA*

## Score = -12 , Var(Score) = 943.3333  
## denominator = 186.9759  
## tau = -0.0642, 2-sided pvalue =0.72023

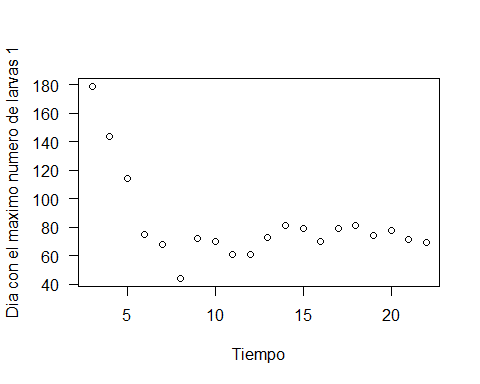


Tendencia casi nula pero esto no es significativo

**TENDENCIA EN L1**

*DIA DE PICO MAXIMO*

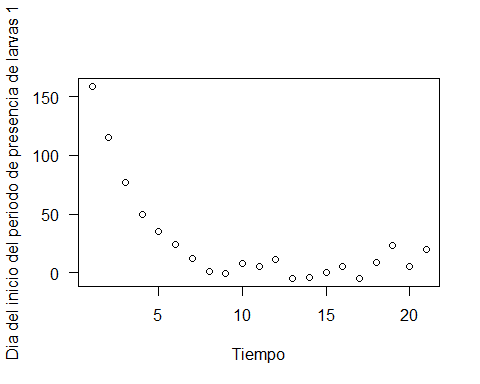
## Score = 12 , Var(Score) = 946  
## denominator = 187.9894  
## tau = 0.0638, 2-sided pvalue =0.72061



Tendencia casi nula pero no es significativo.

*PRIMER DIA*

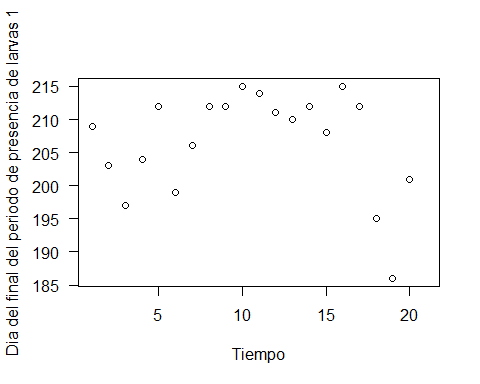
## Score = -92 , Var(Score) = 1092  
## denominator = 207.9904  
## tau = -0.442, 2-sided pvalue =0.005891



Tendencia negativa y significativa. Años de calentamiento!

*ULTIMO DIA*

## Score = 7 , Var(Score) = 932.3333  
## denominator = 184.418  
## tau = 0.038, 2-sided pvalue =0.84422

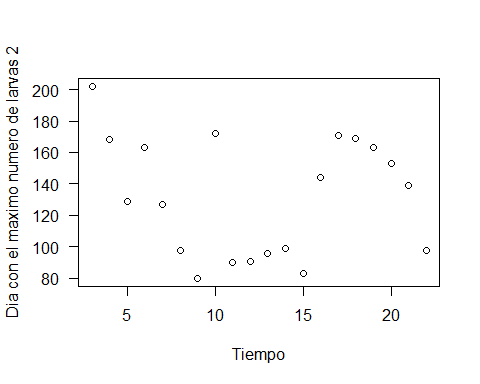


El valor de la intensidad de la tendencia es bajo pero este dato no es significativo.

**TENDENCIA EN L2**

*DIA DE PICO MAXIMO*

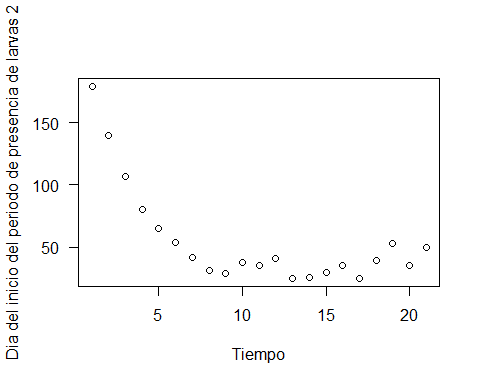
## Score = -2 , Var(Score) = 948  
## denominator = 188.9974  
## tau = -0.0106, 2-sided pvalue =0.97409



El valor de tau es cercano a 0 pero no es significativo; los datos son muy dispares.

*PRIMER DIA*

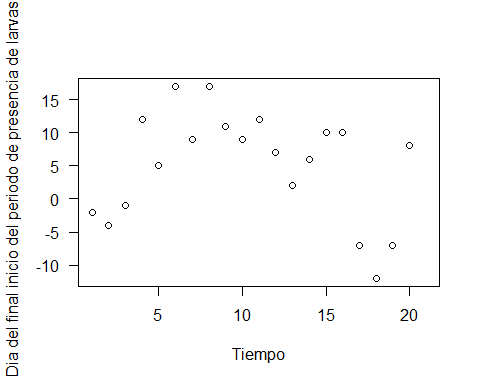
## Score = -92 , Var(Score) = 1092  
## denominator = 207.9904  
## tau = -0.442, 2-sided pvalue =0.005891



Tendencia negativa y significativa debido al calentamiento del modelo.

*ULTIMO DIA*

## Score = -25 , Var(Score) = 945  
## denominator = 187.4833  
## tau = -0.133, 2-sided pvalue =0.43497



Datos muy dispares

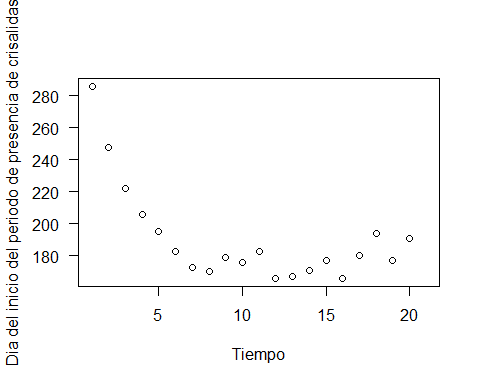
**TENDENCIA EN CRISALIDA**

*DIA DE PICO MAXIMO*

Tendencia debido al calentamiento.

*PRIMER DIA*

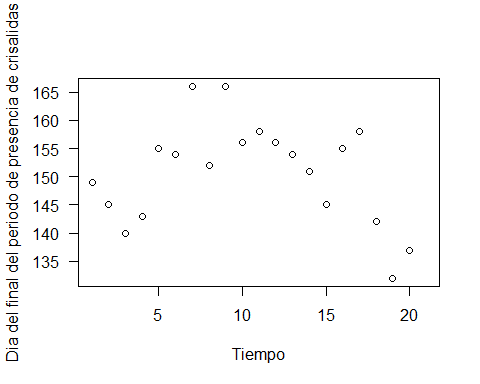
## Score = -69 , Var(Score) = 947  
## denominator = 188.494  
## tau = -0.366, 2-sided pvalue =0.027126



Tendencia negativa significativa. Calentamiento

*ULTIMO DIA*

## Score = -18 , Var(Score) = 944  
## denominator = 186.9759  
## tau = -0.0963, 2-sided pvalue =0.58006

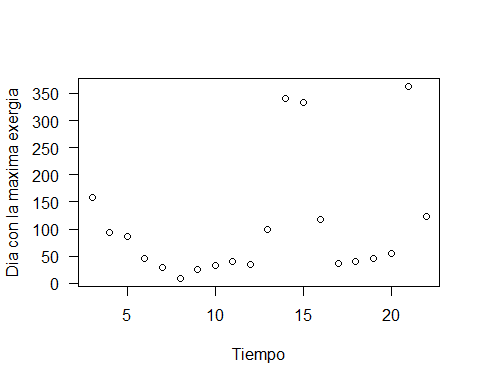


Valor de la tendencia cercano a cero, pero no significativo. Datos muy dispares.

**TENDENCIA EN EXERGIA**

*DIA DLE PICO MAXIMO*

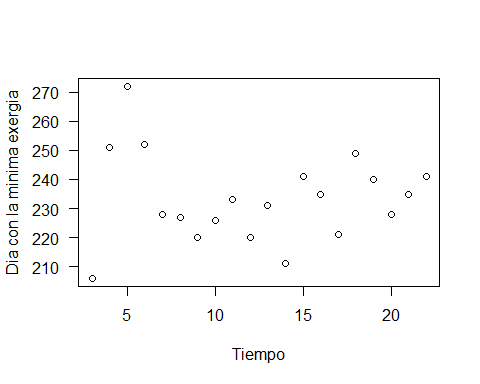
## Score = -23 , Var(Score) = 949  
## denominator = 189.4993  
## tau = -0.121, 2-sided pvalue =0.47513



Datos muy dispares en algunos años, por eso no sale significativo.

*DIA DEL PICO MINIMO*

## Score = 28 , Var(Score) = 946  
## denominator = 187.9894  
## tau = 0.149, 2-sided pvalue =0.38003



Calentamiento del modelo.