

# DADES EN RETICULES

## Introduction to WinBUGS and GeoBUGS

Dra. ROSA ABELLANA

DEPARTAMENT SALUT PÚBLICA. UNIVERSITAT DE  
BARCELONA

# BUGS program

## Bayesian inference using Gibbs Sampling

- Language for specifying complex Bayesian models
- Constructs object-oriented internal representation of the model
- Simulation from full conditionals using Gibbs sampling
- Current version (WinBUGS 1.4) runs in Windows
- 'Classic' BUGS available for UNIX but this is an old version

WinBUGS is freely available from

**<http://www.mrc-bsu.cam.ac.uk/bugs>**

- Scripts enable WinBUGS 1.4 to run in batch mode or be called from other software
- Interfaces developed for R, Excel, Splus, SAS, Matlab
- OpenBUGS site <http://www.rni.helsinki.fi/openbugs> provides an open source version2-

# WinBUGS formats of data

Type of formats :

## 1. Rectangular

```
n[ ]  r[ ]  
47    0  
148  18  
...  
360  24  
END
```

## 2. S-Plus:

```
list(N=12,n = c(47,148,119,810,211,196,  
148,215,207,97,256,360),  
r = c(0,18,8,46,8,13,9,31,14,8,29,24))
```

## Function BUGS

- $p \leftarrow \text{step}(x-0.7) = 1$  if  $x \geq 0.7$ ; 0 otherwise.

E.g: Average the samples of  $p$ , obtain the probability that  $x \geq 0.7$

- $p \leftarrow \text{equals}(x,.7) = 1$  si  $x = 0.7$ , 0 otherwise.
- $\text{tau} \leftarrow 1/\text{pow}(s,2)$  compute  $= 1/s^2$ .
- $s \leftarrow 1/\text{sqrt}(\text{tau})$  compute  $s = 1/\tau^{1/2}$ .

- Look 'Model Specification/Logical nodes' of the winBUGS manual for all the functions.

## Most common distributions

Expression	Distribution	Us
dbin	binomial	$r \sim \text{dbin}(p, n)$
dnorm	normal	$x \sim \text{dnorm}(\mu, \tau)$
dpois	Poisson	$r \sim \text{dpois}(\lambda)$
dunif	uniform	$x \sim \text{dunif}(a, b)$
dgamma	gamma	$x \sim \text{dgamma}(a, b)$

Normal distribution is parameterized by the mean and the precision (inverse of variance) =  $1/\text{sd}^2$ .

Look 'Model Specification/The BUGS language: stochastic nodes/Distributions'

.

## Some aspects of language BUGS

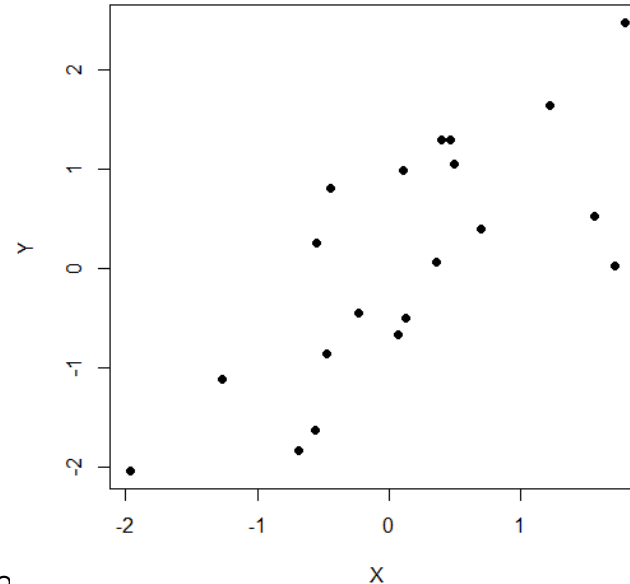
- $\sim$  represents stochastic dependence, e.g.  $r \sim \text{dunif}(a,b)$
- Can use arrays and loops

```
for (i in 1:n){  
  r[i] ~ dbin(p[i],n[i])  
  p[i] ~ dunif(0,1)  
}
```
- Some functions can appear on left-hand-side of an expression, e.g.

```
logit(p[i])<- a + b*x[i]  
log(m[i]) <- c + d*y[i]
```
- $\text{mean}(p[])$  to take mean of whole array,  $\text{mean}(p[m:n])$  to take mean of elements  $m$  to  $n$ . Also for  $\text{sum}(p[])$ .

# Simple linear regression in WinBUGS

We observed the data:



Call:

```
lm(formula = Y ~ X)
```

Residuals:

Min	1Q	Median	3Q	
-1.5123	-0.6019	-0.1857	0.8309	1.2

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.04017	0.19197	-0.209	0.836588
X	0.92174	0.20027	4.603	0.000221 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8491 on 18 degrees of freedom

Multiple R-squared: 0.5406, Adjusted R-squared: 0.5151

F-statistic: 21.18 on 1 and 18 DF, p-value: 0.000221

## WinBUGS:

### #The model

```
model{
  for(i in 1:n){
    y[i]~dnorm(mu[i],tau)  #likelihood distribution
    mu[i] <- alpha + beta*x[i]
  }
  #prior distributions
  tau~dgamma(0.01,0.01)
  alpha~dnorm(0,0.01)
  beta~dnorm(0,0.01)
}
```

### #The data

```
list(n=20,
X=c(-0.5605,-0.2302,1.5587,0.0705,0.1293,1.7151,0.4609,-1.2651,-0.6869,-
0.4457,1.2241,0.3598,0.4008,0.1107,-0.5558,1.7869,0.4979,-1.9666,0.7014,-
0.4728),
Y=c( -1.6283,-0.4482,0.5327,-0.6584,-0.4958,0.0284,1.2987,-1.1117,-
1.825,0.8082,1.6505,0.0647,1.2959,0.9888,0.2657,2.4756,1.0518,-
2.0285,0.3954,-0.8533))
```



## #The initial values

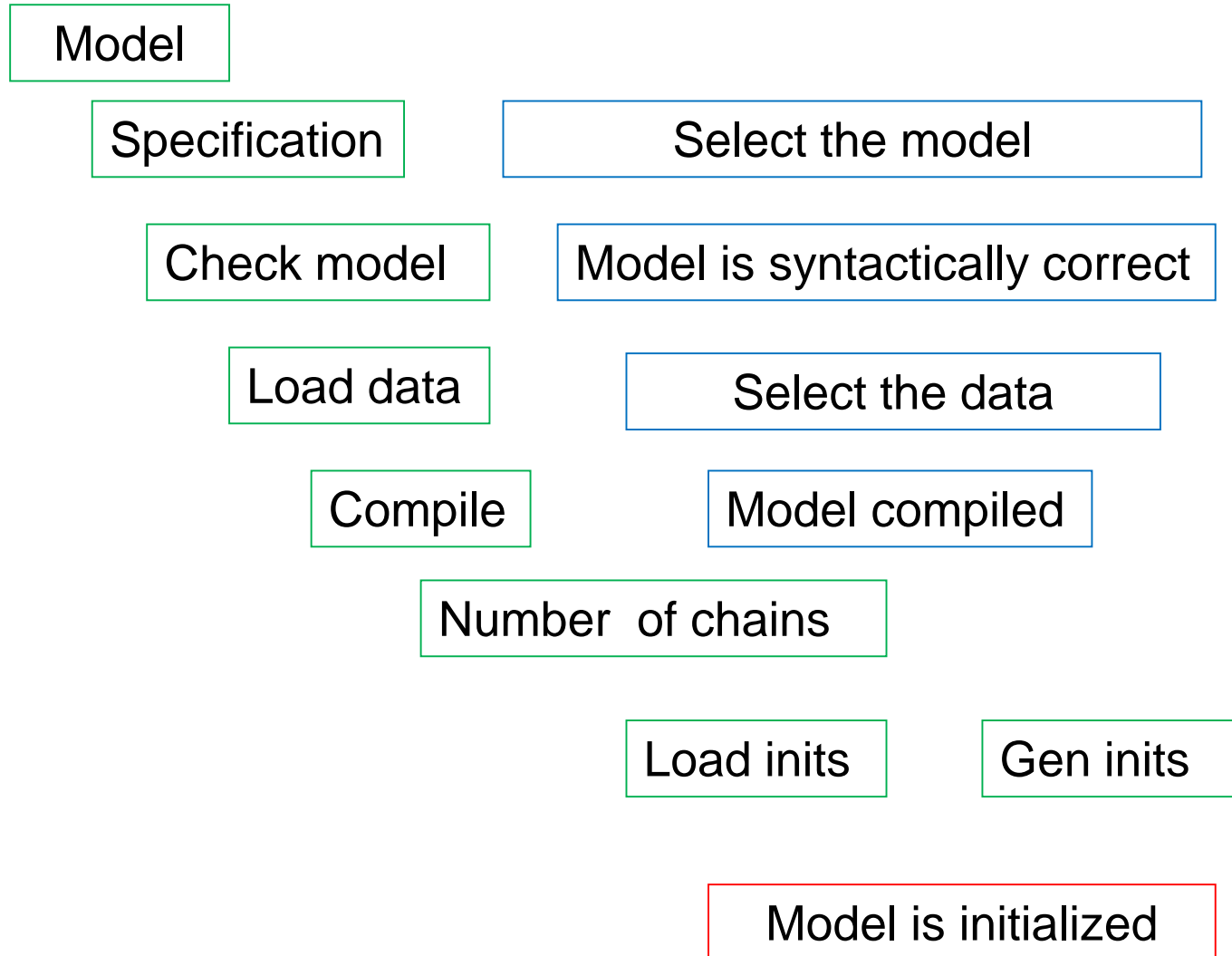
Chain 1:

```
list(alpha=1, beta=0, tau=0.01)
```

Chain 2:

```
list(alpha=-1,beta=-2,tau=0.8)
```

## 1. Running WinBUGS:



2. For each variable you'd like to analyze, go to  
**"inference/sample"**

Enter the name of the variable in the "nodes" box,  
and click "set". Samples of variables you don't  
request to record are automatically discarded.

3. Check the convergence of the chains

**Menu Inference/Trace plot**

**Menu Inference/bgr diag**

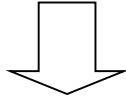
**Menu inference/Autocorrelation function**

4. Goodness of fit, when the samples had already converge

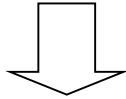
**Menu Inference/DIC**

Sampling of Gibbs       $\longrightarrow$       We stroll for the sample space

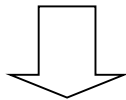
First samples are discarded



Burning process



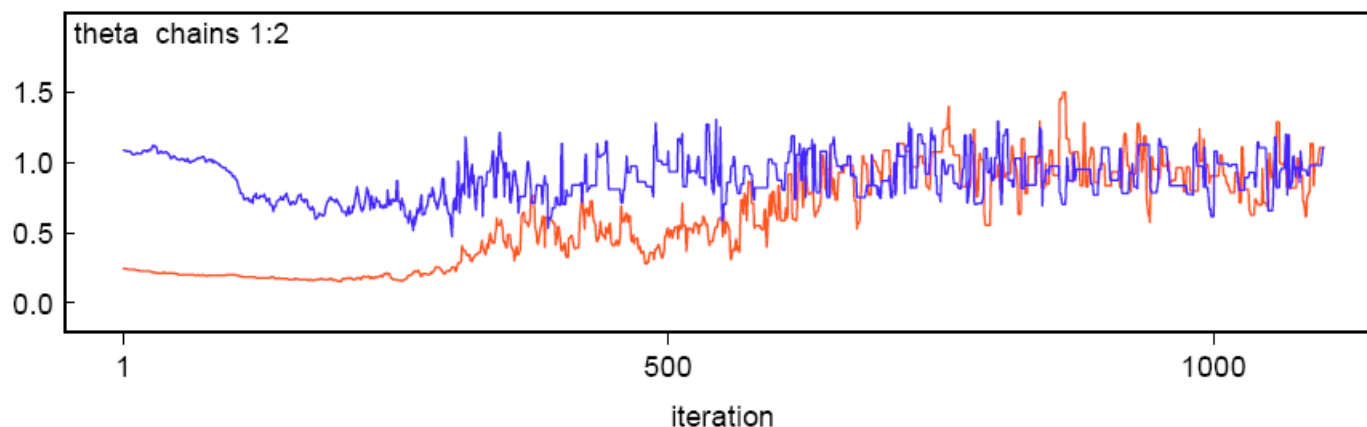
Check if the chains came from posterior  
distribution



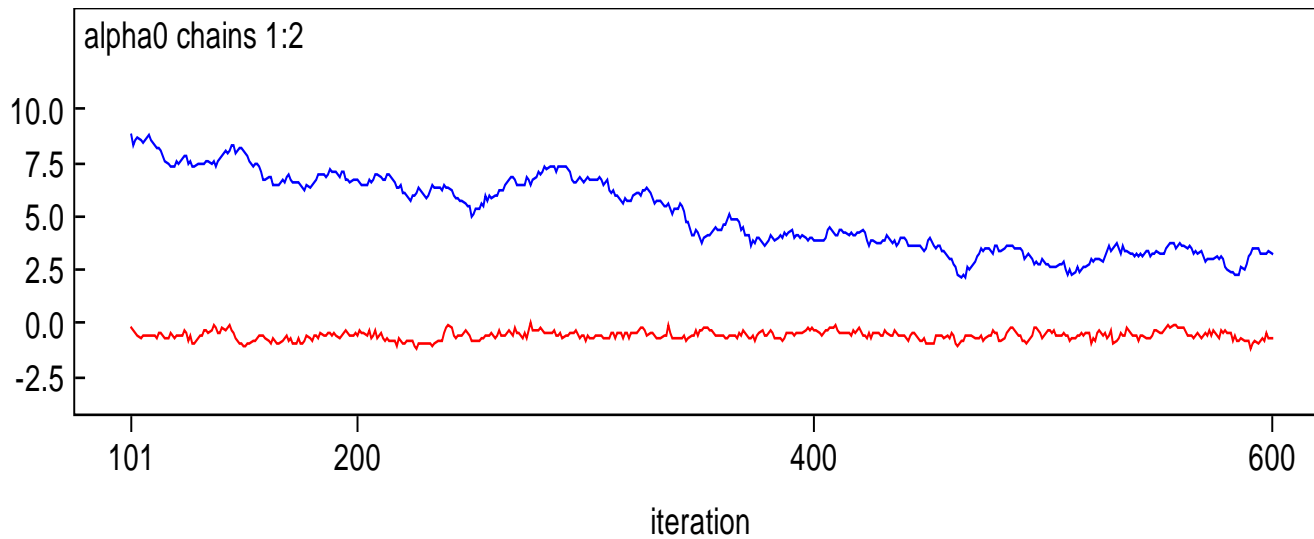
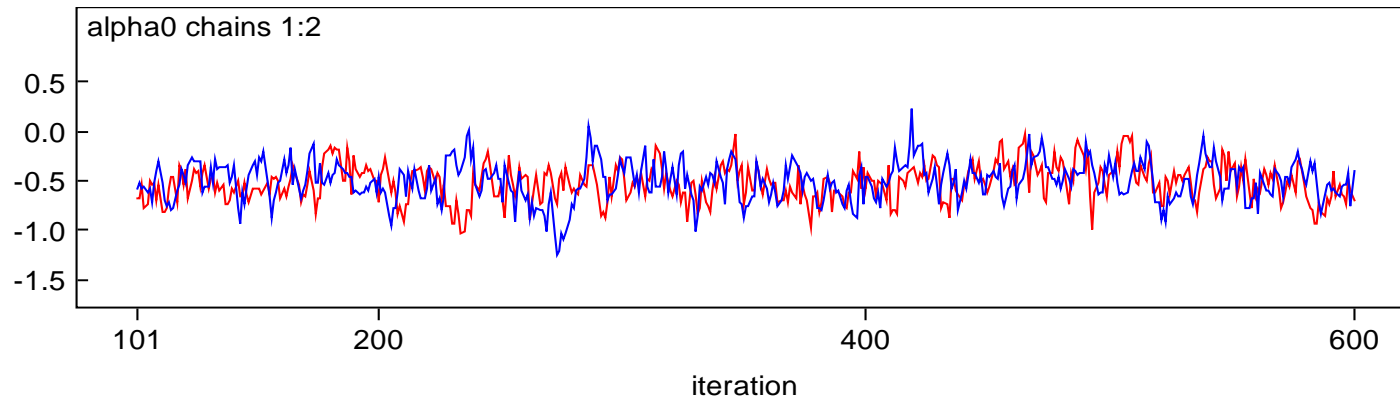
*Convergence* to posterior distribution.

## Checking convergence

- Convergence is to a distribution not to a single value.
- Once convergence reached, samples should look like a random scatter about a stable mean value.
- Practice, is to *run* multiple chains for the same model but with different initial points, and visually inspect “trace” plot for convergence



## Example convergence and no convergence



## Statistician of Gelman-Rubin

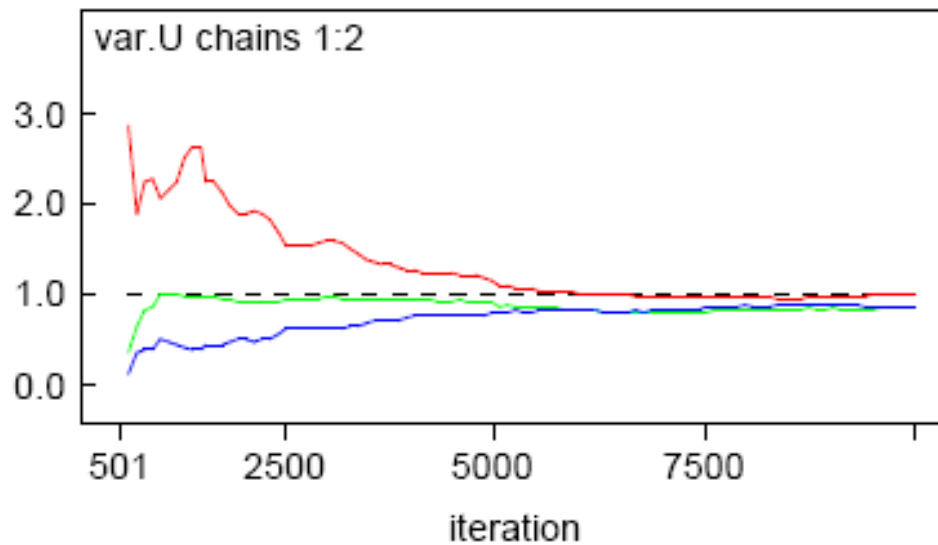
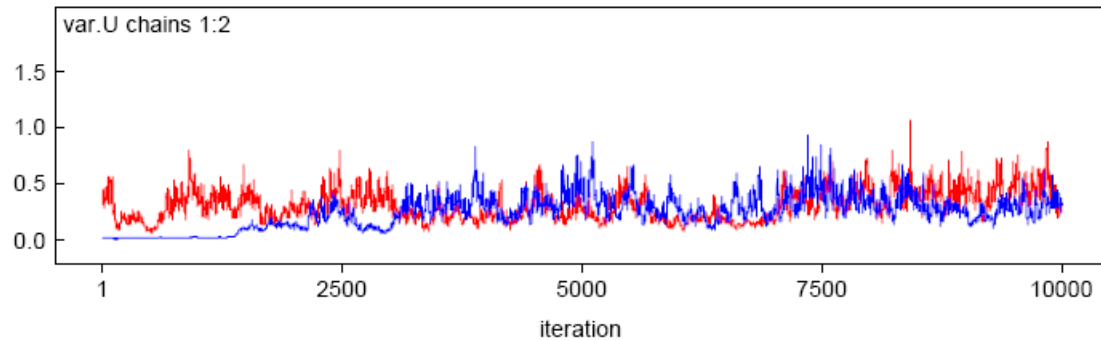
Formally, can use the Gelman-Rubin statistic based in the ratio of the *between* and *within* chains variance (ANOVA)

Gelman: the best to identify the no convergence, starting from different initial values,

Intuitively, the behavior of all the chains would have to be the same.

Or, the variance within the chain would have to **be the same** that the variance along the chain

## Example of Gelman and Rubin with the Winbugs



*B pooled*

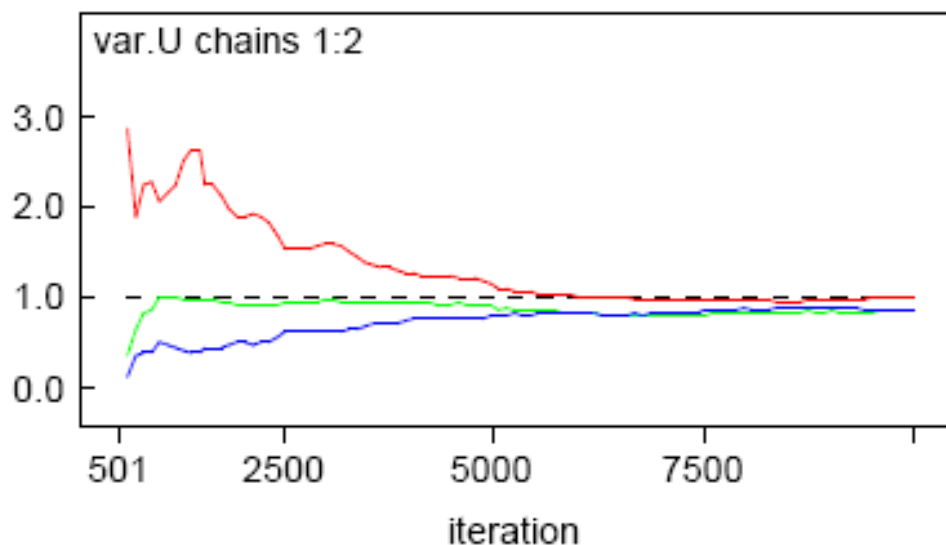
*W average of width*



## Gelman-Rubin statistic

- Generate multiple chains starting at *over-dispersed* initial values.
- Denote the number of chains generated by  $M$  and the length of each chain by  $2T$ .
- Take as a measure of posterior variability the width of the  $100(1 - \alpha)\%$  credible interval for the parameter of interest (in WinBUGS,  $\alpha = 0.2$ ).
- From the final  $T$  iterations we calculate the empirical credible interval for each chain. We then calculate the average width of the intervals across the  $M$  chains and denote this by  $W$
- Calculate the width  $B$  of the empirical credible interval based on all  $MT$  samples pooled together.

The ratio  $R = B / W$  of pooled to average interval widths should be greater than 1 if the starting values are suitably overdispersed; it will also tend to 1 as convergence is approached, and so we might assume convergence for practical purposes if  $R < 1.05$ , say.



# GeoBUGS

GeoBUGS is an add-on module to WinBUGS which provides an interface for:

- producing maps of the output from disease mapping and other spatial models

- \* creating and manipulating adjacency matrices that are required as input for the conditional autoregressive models available in WinBUGS 1.4 for carrying out spatial smoothing.

Version 1.2 of GeoBUGS contains map files for

- \* Districts in Scotland (called **Scotland**)
- \* Wards in a London Health Authority (called **London\_HA**)
- \* Counties in Great Britain (called **GB\_Counties**)
- \* Departements in France (called **France**)
- \* Nomoi in Greece (called **Greecenomoi**)
- \* Districts in Belgium (called **Belgium**)
- \* Communes in Sardinia (called **Sardinia**)
- \* Subquarters in Munich (called **Munich**)
- \* A 15 x 15 regular grid (called **Elevation**)
- \* Wards in West Yorkshire (UK) (called **WestYorkshire**)
- \* A 4 x 4 regular grid (called **Forest**)
- \* A grid of 750 m<sup>2</sup> grid cells covering the town of Huddersfield and surroundings in northern England (called **Huddersfield\_750m\_grid**)

A list of the area IDs for each map and the order in which the areas are stored in the map file can be obtained using the **export Splus** command.

GeoBUGS 1.2 also has facilities for importing user-defined maps reading polygon formats from Splus, ArcInfo and Epimap, plus a link to a program written by Yue Cui for importing ArcView shape files.

<http://www.biostat.umn.edu/~yuecui/>

## Export S-plus map to Winbugs

**sp2WB** (library maptools)

The function exports an sp SpatialPolygons object into a S-Plus map format to be import by Win-BUGS.

## In Winbugs

Open the polygon file as a separate text file in WinBUGS 1.4Beta and select the appropriate **import** option from the ***Map menu.***