

Clasificación de Estrellas Variables en CU7

Monthly Notices
of the

ROYAL ASTRONOMICAL SOCIETY



Mon. Not. R. Astron. Soc. (2011)

doi:10.1111/j.1365-2966.2011.18575.x

Random forest automated supervised classification of *Hipparcos* periodic variable stars

P. Dubath,^{1,2*} L. Rimoldini,^{1,2} M. Süveges,^{1,2} J. Blomme,³ M. López,⁴ L. M. Sarro,⁵
J. De Ridder,³ J. Cuypers,⁶ L. Guy,^{1,2} I. Lecoœur,^{1,2} K. Nienartowicz,^{1,2} A. Jan,^{1,2}
M. Beck,^{1,2} N. Mowlavi,^{1,2} P. De Cat,⁶ T. Lebzelter⁷ and L. Eyer^{1,2}

Diego Ordóñez Blanco
Observatorio de Ginebra / ISDC
Gaia CU7 team
diego.ordonez@unige.ch

Septiembre, 2011

Contexto

- Progreso del trabajo del equipo de CU7 para preparar el análisis de datos de Gaia
- Se ha empleado en este trabajo el catalogo de estrellas de Hipparcos como ejemplo
 - Determinar la mejor estrategia de clasificación
- Elección actual: Clasificación en tres etapas.

Sample of Surveyed Stars

Variable Stars (10% to 30%)

Non-Periodic Stars

L

SR

I

Be

GCAS

RS

BY

Periodic Stars

SPB

DSCT

ECL

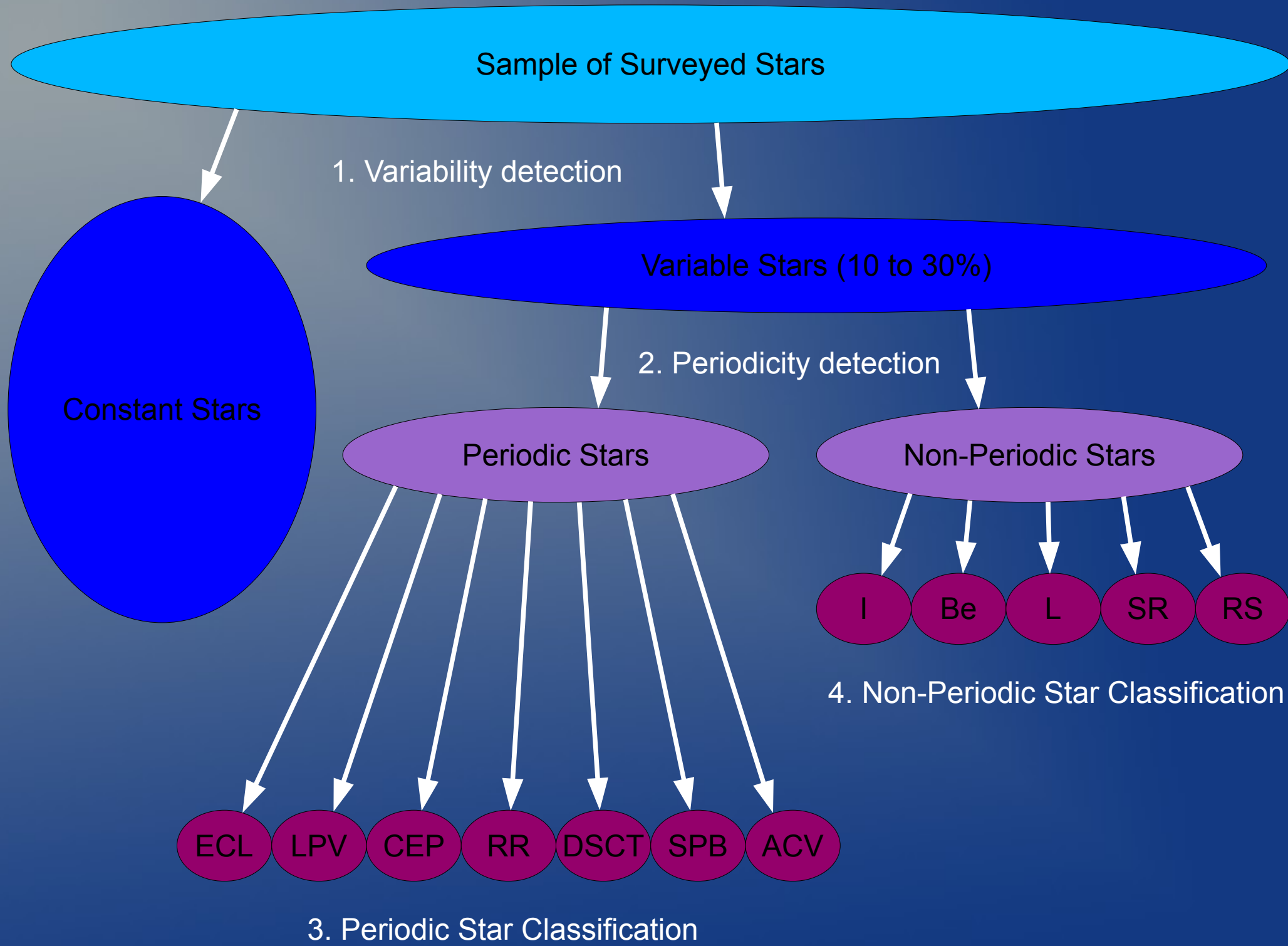
CEP

ACV

LPV

RR

Constant Stars



Sample of Surveyed Stars

1. Variability detection

Variable Stars (10 to 30%)

2. Periodicity detection

Constant Stars

Periodic Stars

Non-Periodic Stars

ECL

LPV

CEP

RR

DSCT

SPB

ACV

3. Periodic Star Classification

I

Be

L

SR

RS

4. Non-Periodic Star Classification

Detección de Variabilidad

- Criterios

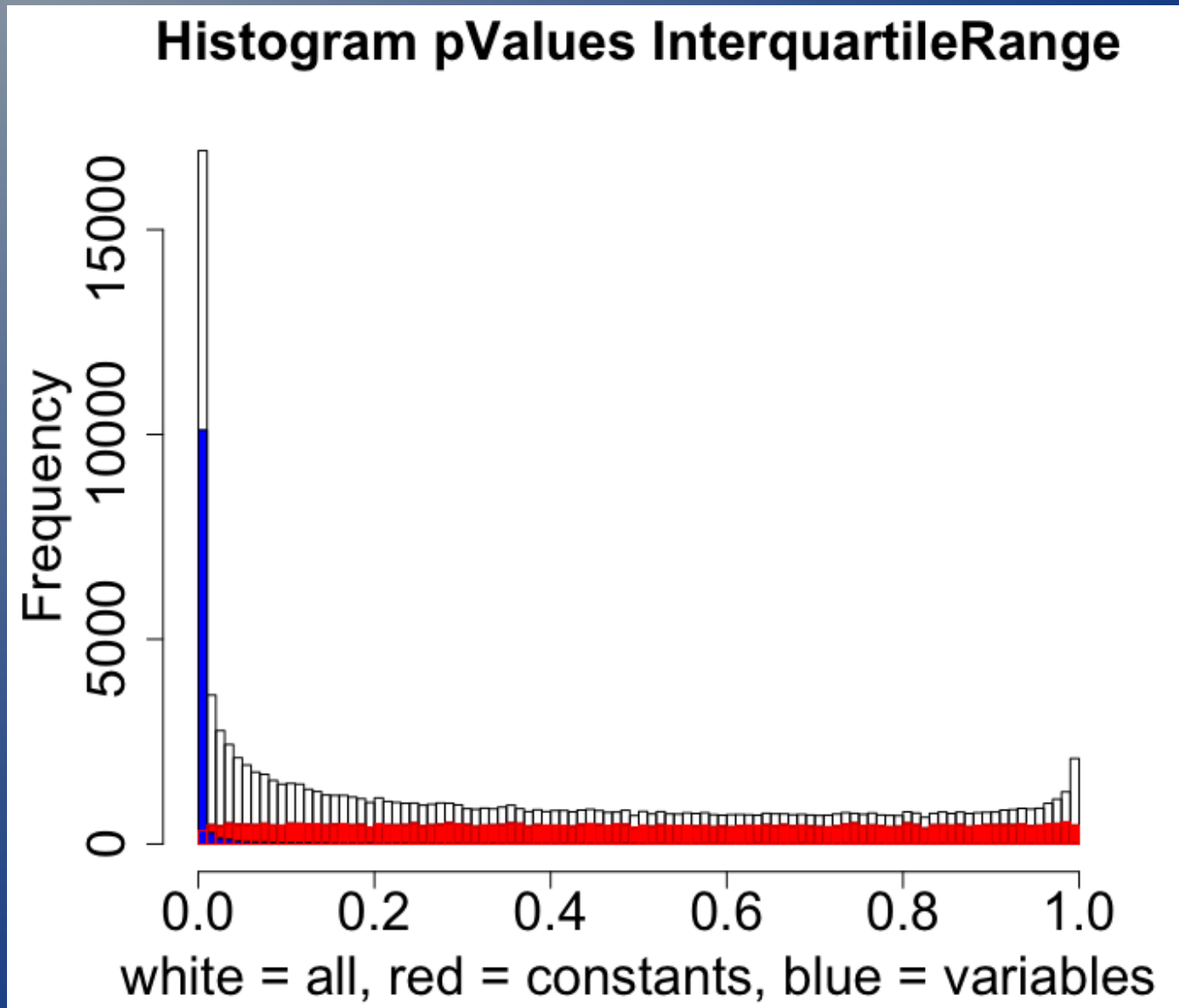
Stetson	IQR
Chi squared	Abbe

$$\chi^2 = \sum_{i=1}^N \left(\frac{x_i - \bar{x}}{\sigma} \right)^2$$

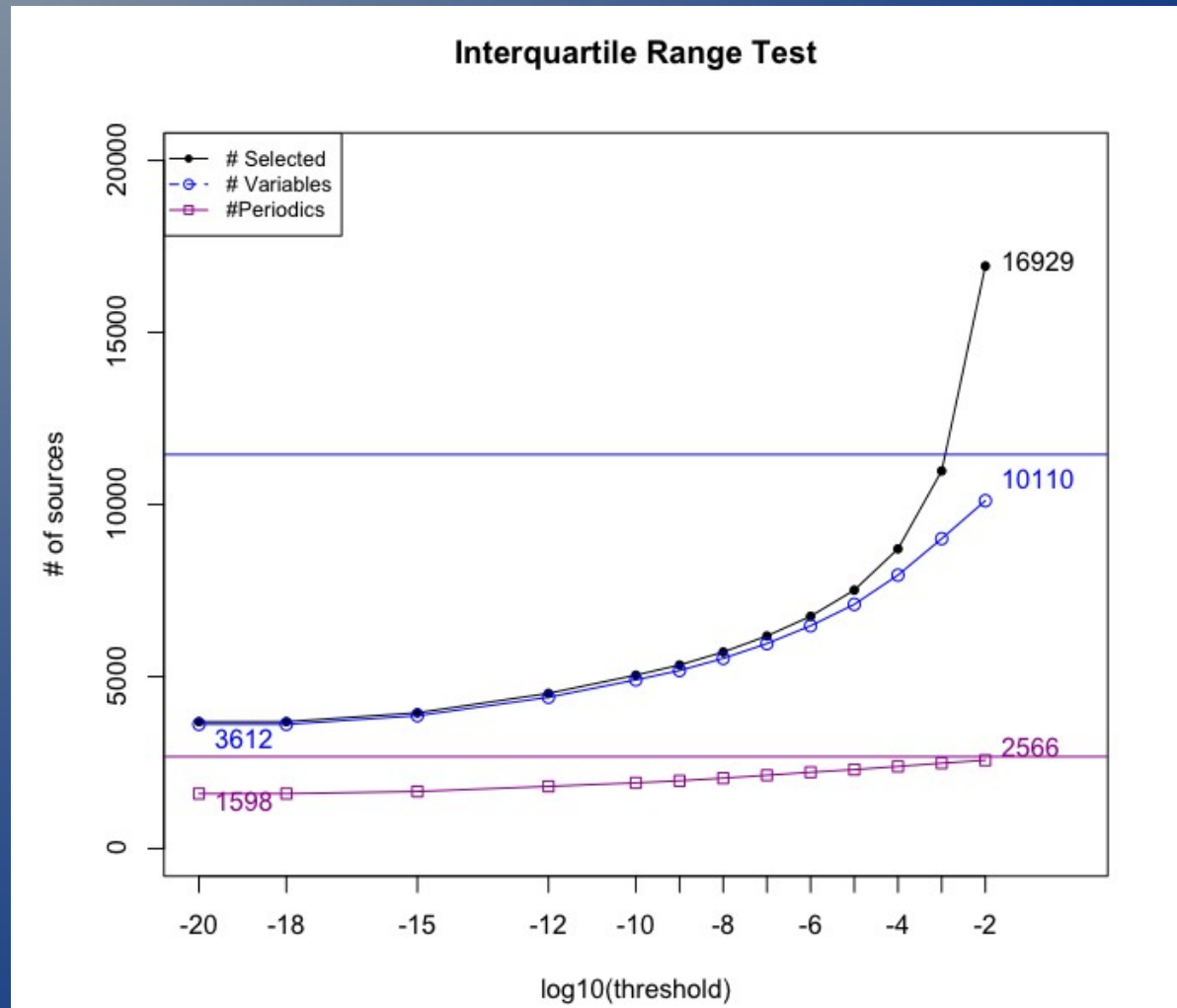
$$Abbe = \frac{1}{2} \frac{\sum_{i=1}^{n-1} (x_{i+1} - x_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

- Calcular los valores p : probabilidad de la hipótesis nula $H_0 =$ estrella constante
 - If pValue < 1e-4 → estrella variable

Hipparcos

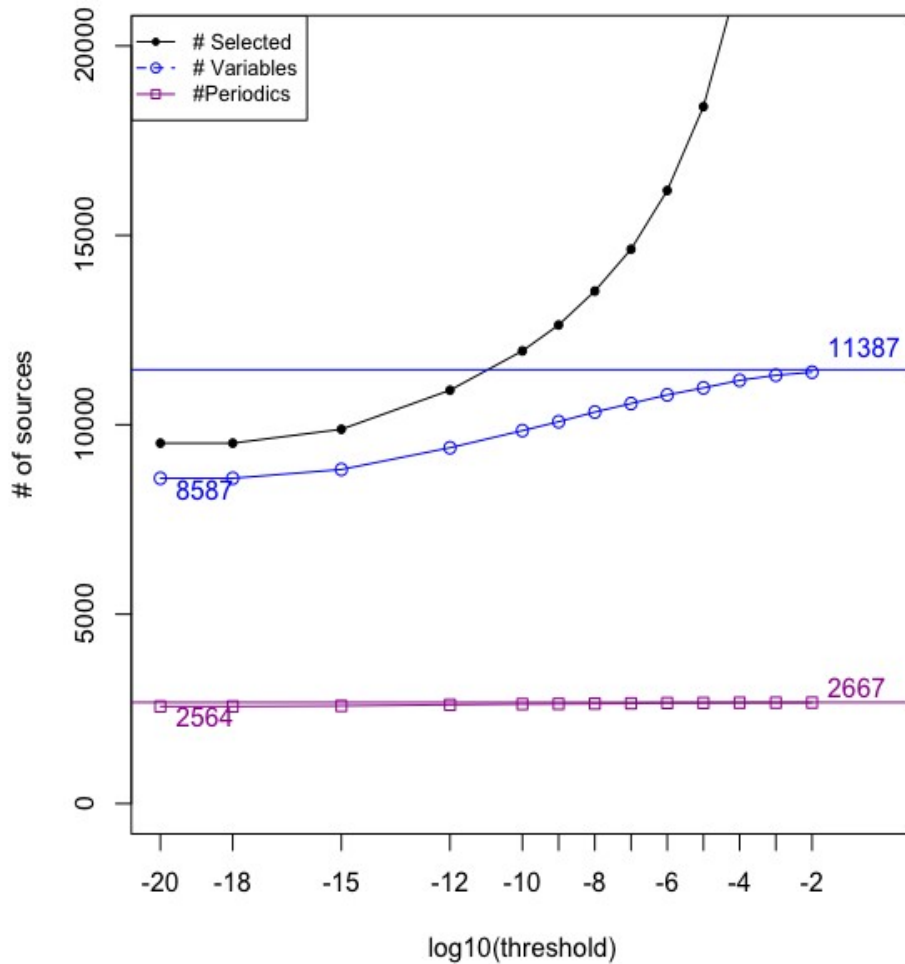


IQR

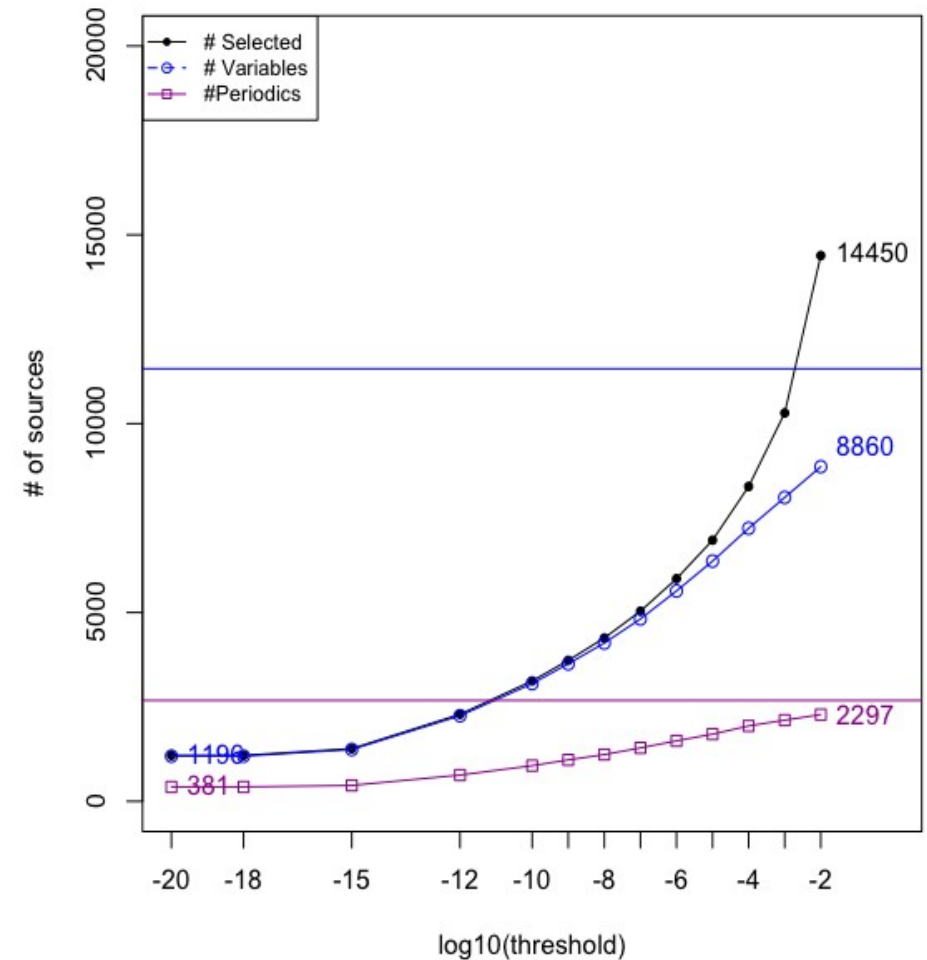


Valor Umbral

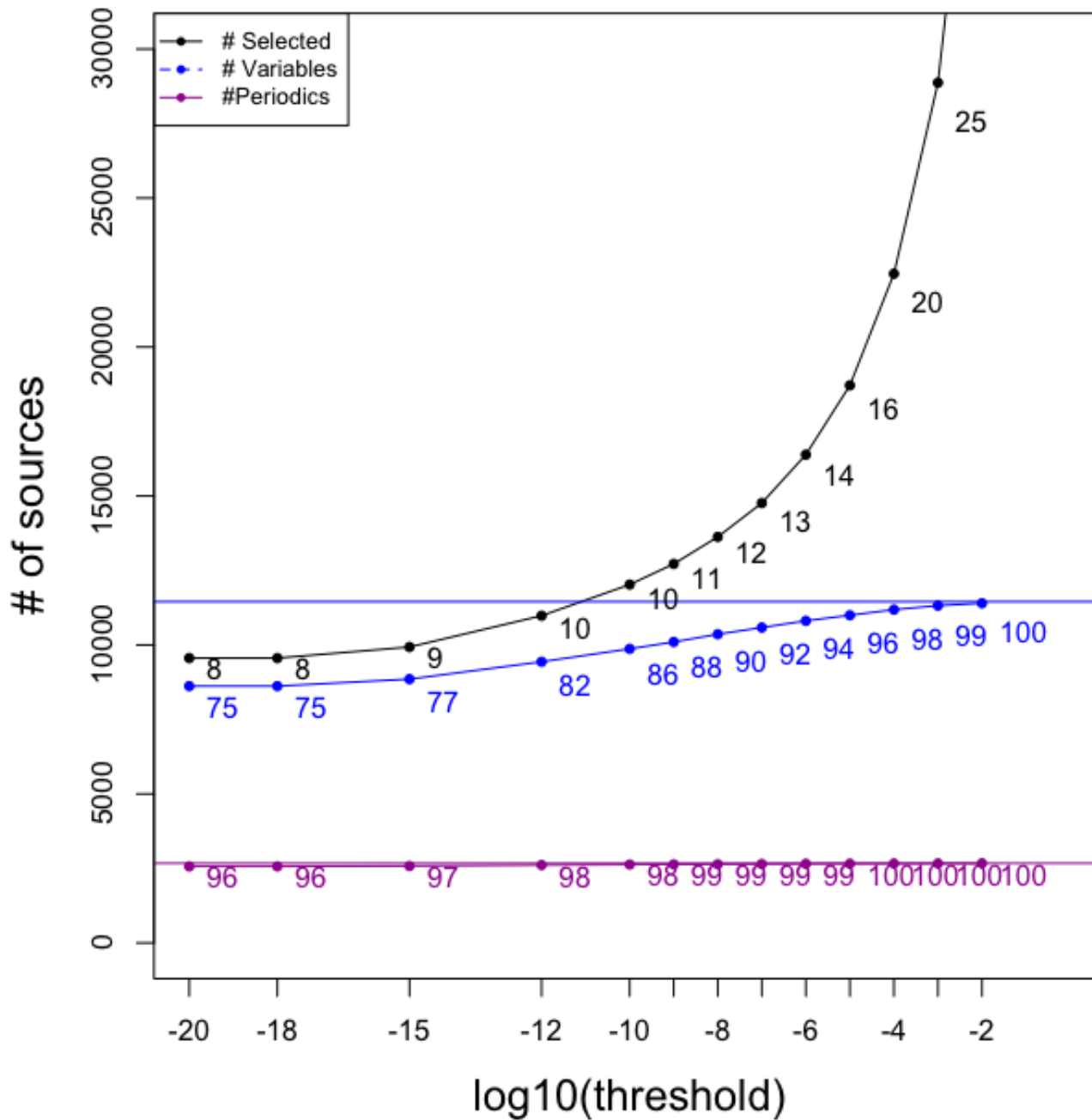
Chi2 Test



Abbe Test

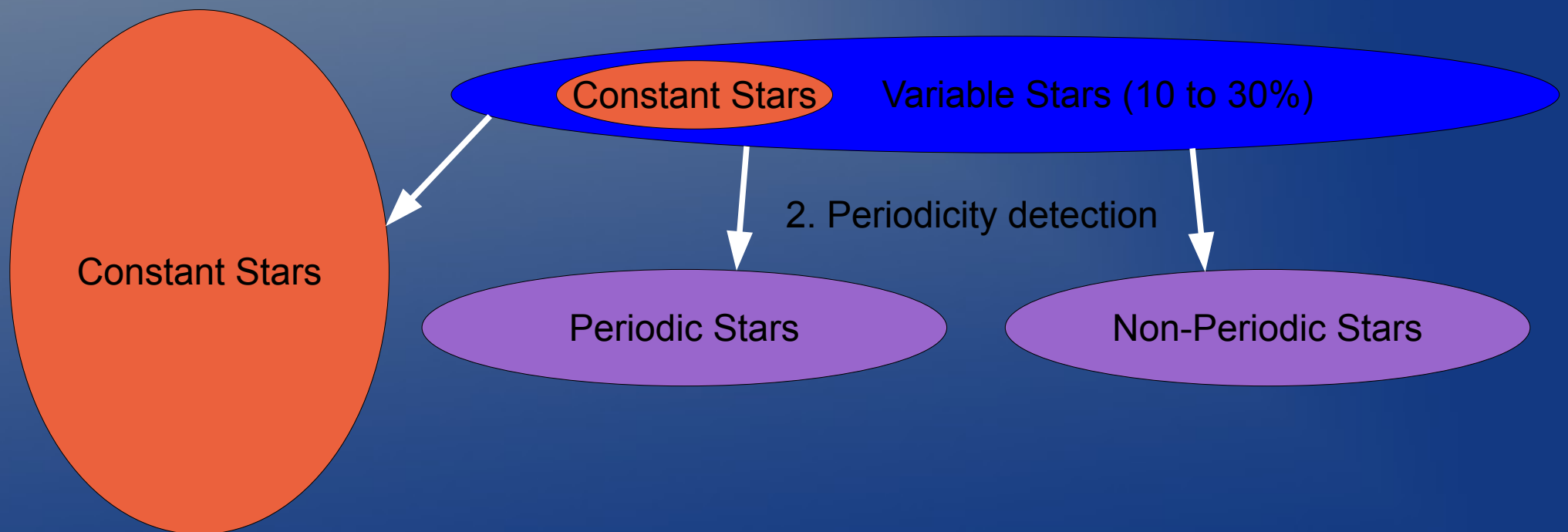


Union (Abbe/Stet/IQR/Chi2/Kur/Out)Test

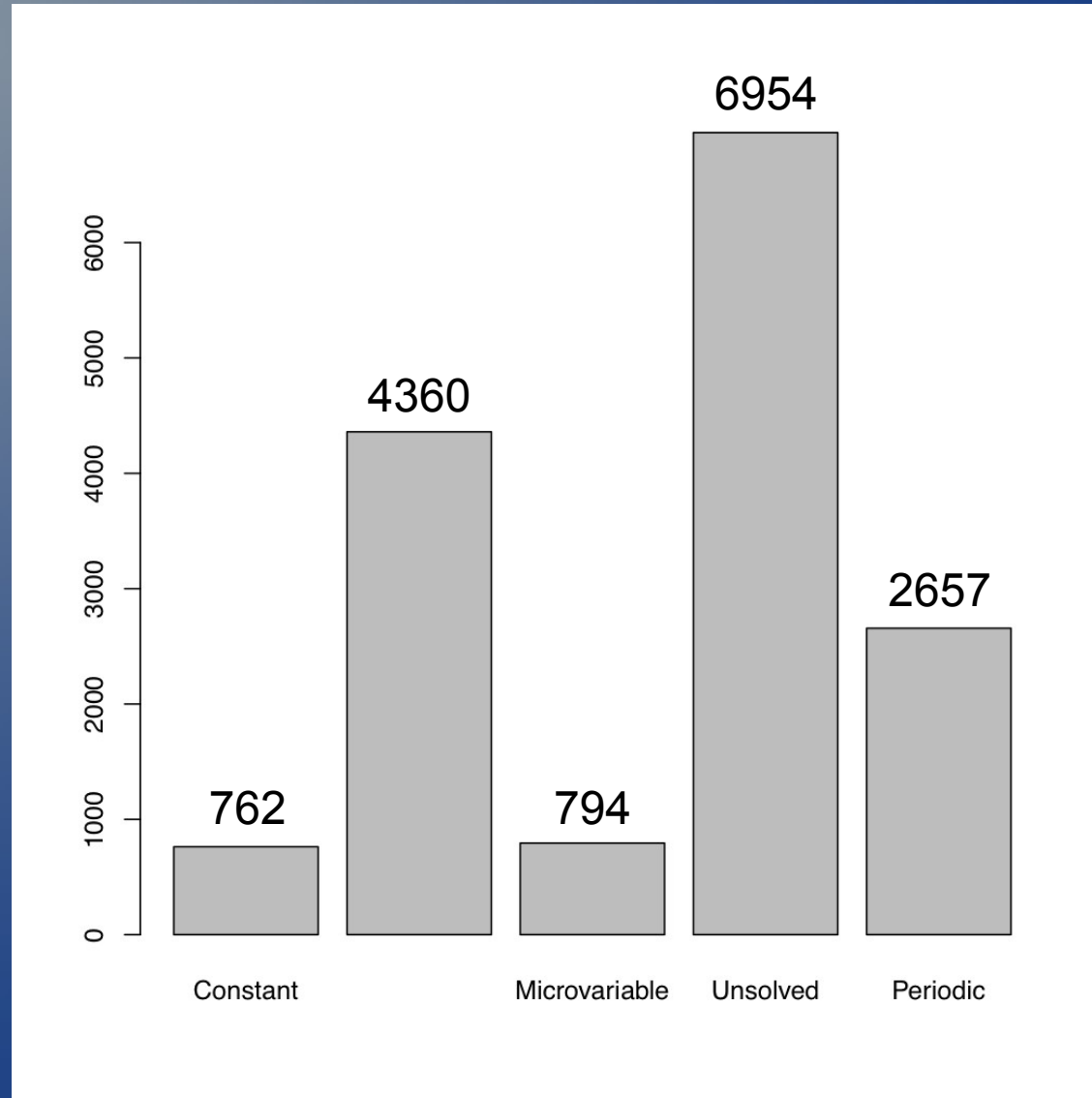


Elección para la detección de Variabilidad

- Stetson ($p\text{Values} < 1e-2$) + IQR ($p\text{Values} < 1e-3$)
 - 17'006 candidatas (14.8 % del total)



Tipos de Variabilidad en Hipparcos

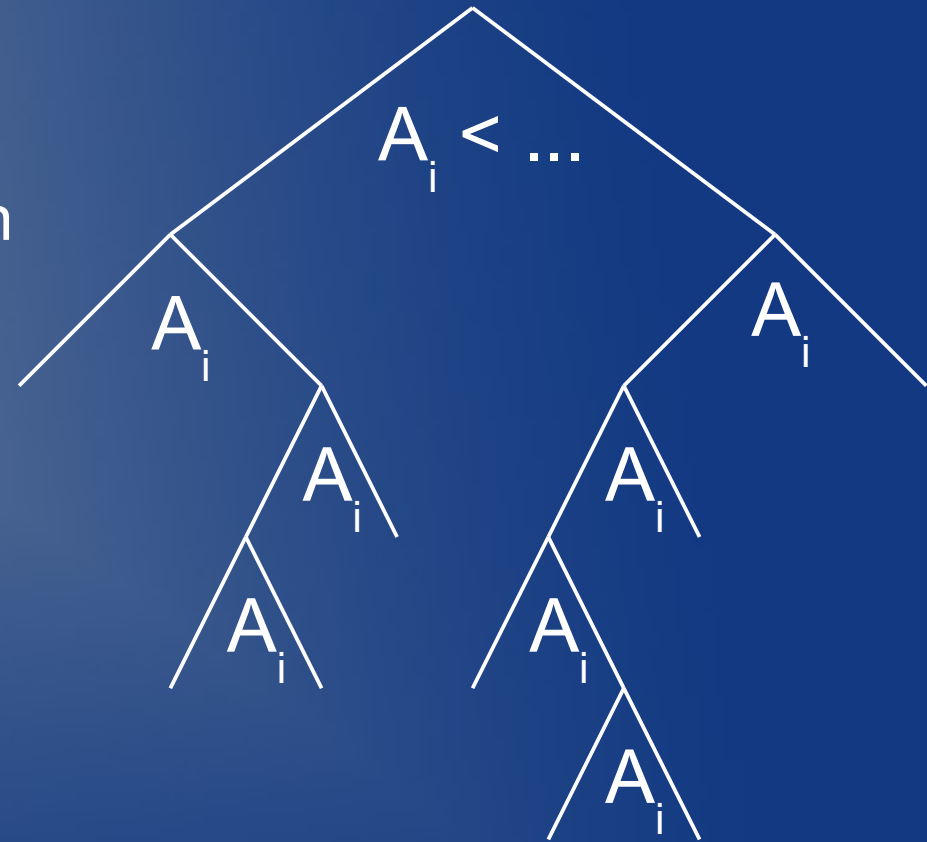


Detección de periodicidad utilizando clasificación supervisada

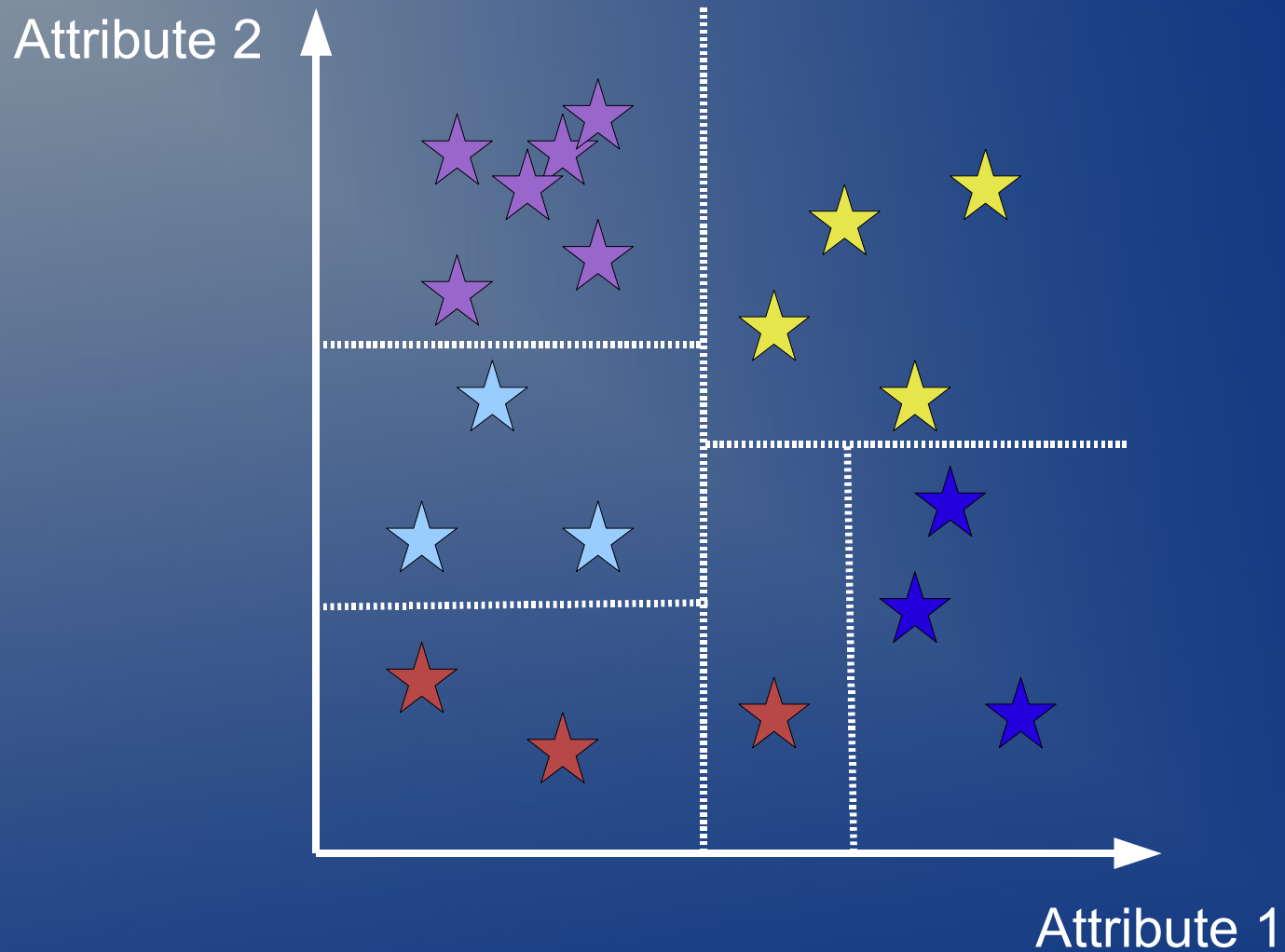
- Calcular atributos que caracterizan la curva de luz
- Utilizar los atributos como variables de entrada en un clasificador supervisado (R. Forest).
- Entrenar el clasificador con un conjunto de estrellas de tipos conocidos.
- Utilizar validación cruzada para evaluar el rendimiento de esta aproximación.

Random Forest (1/3)

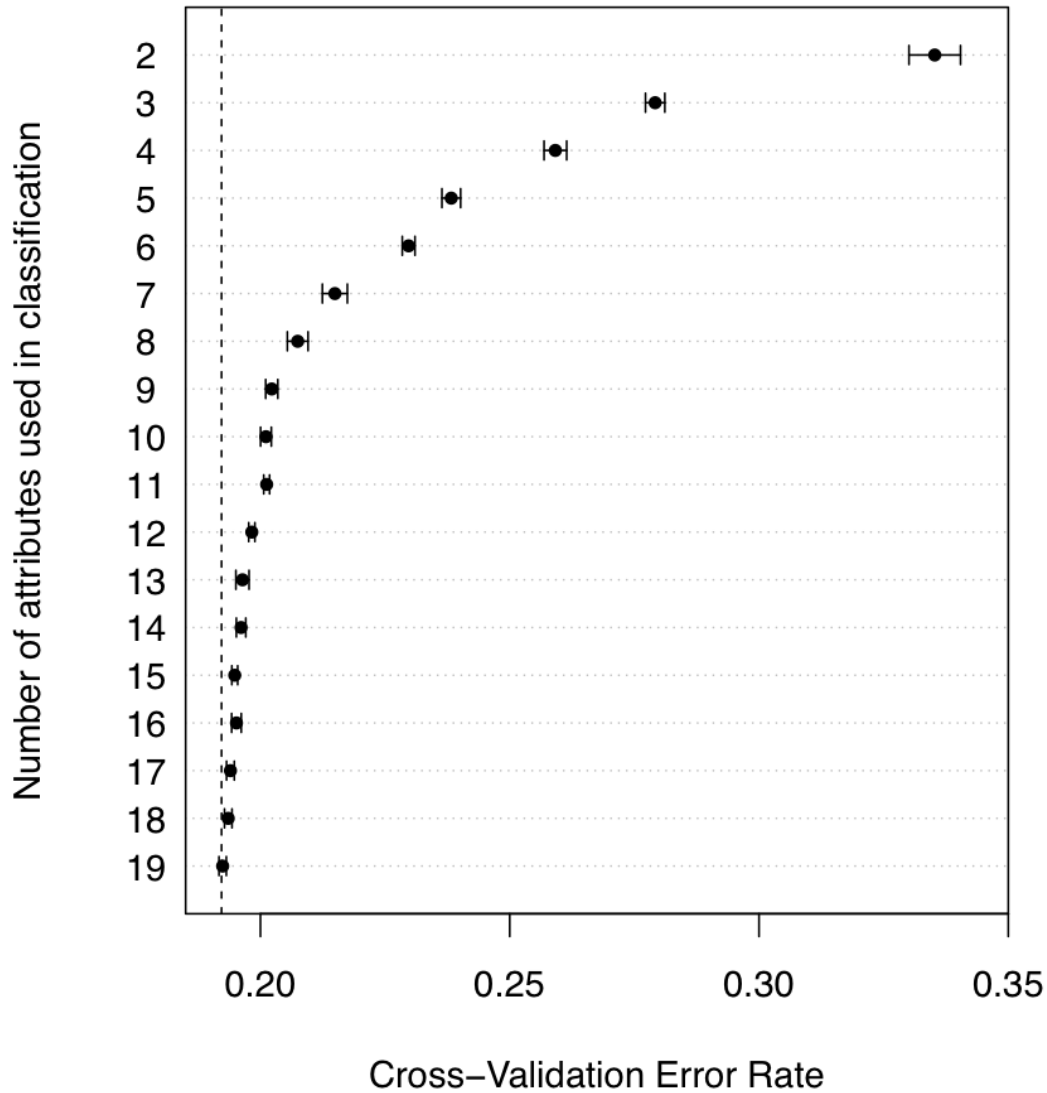
- Árboles de Clasificación
- Particiones binarias
- Menor impureza en cada partición



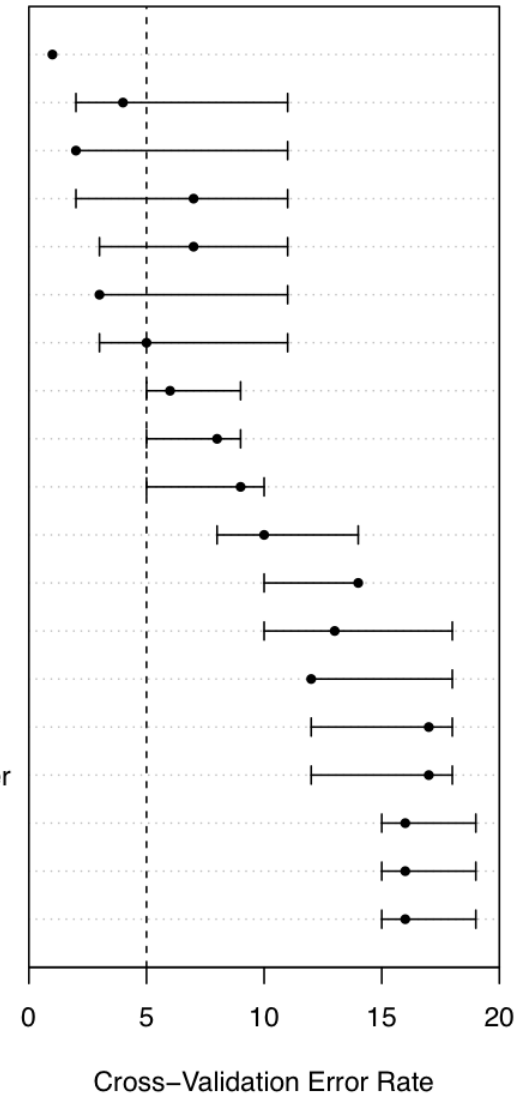
Random Forest (2/3)



Random Forest y CV

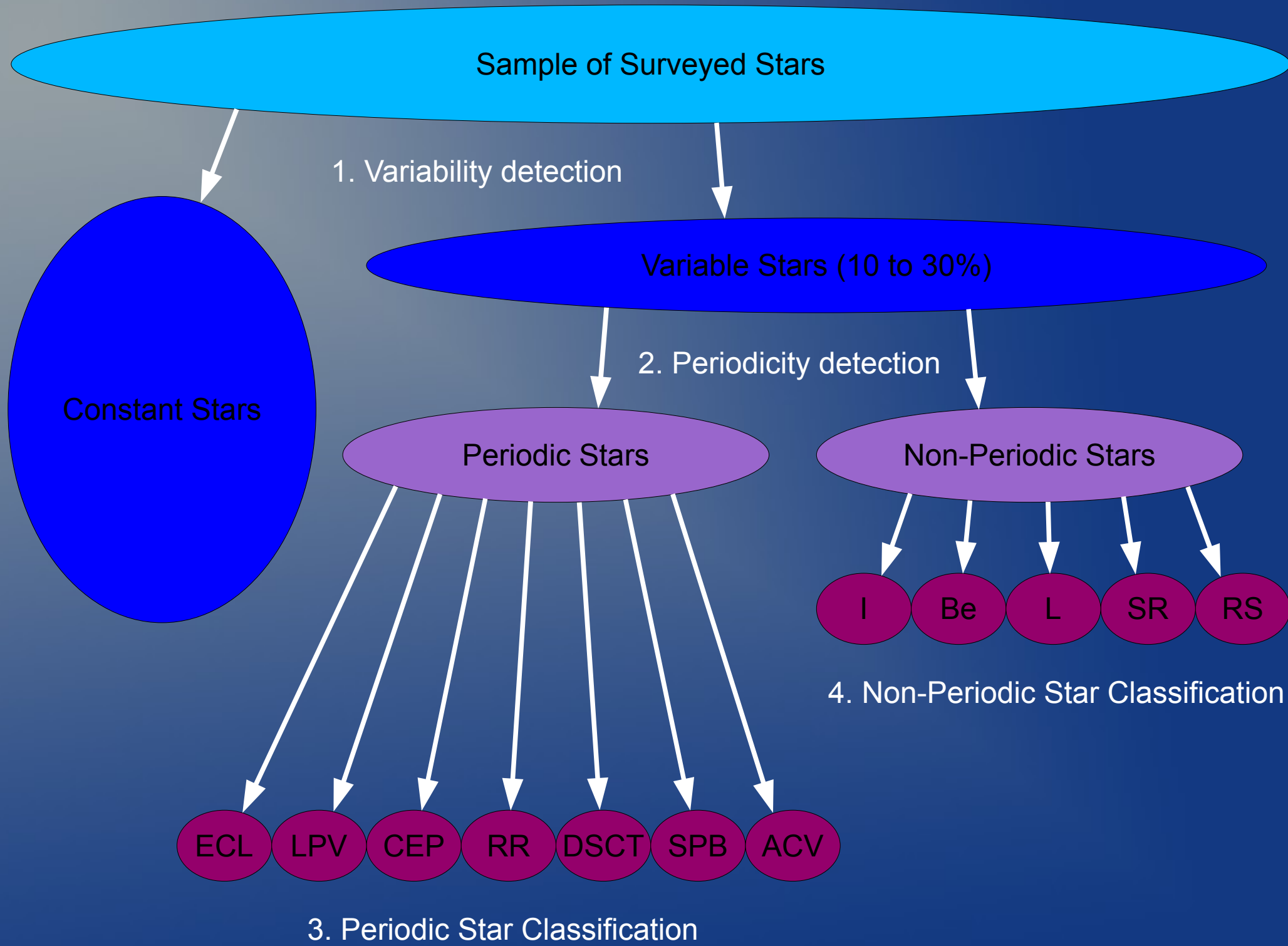


- WstetsonJ
- Raw_LogRange
- p2pScatterNorm2OnFoldedTS
- qsoVar
- logPqso
- FAP0
- p2pScatterFoldedDetrendedTSratio
- variation5Bin2
- Raw_RobustWeightedStdDev
- logPnonQso
- VI_Color
- scatterResidualDetrMeanTSratio
- variation5MedianBin5
- variation5MedianBin2
- WstetsonK
- Res_WeightedNormalizedP2pScatter
- variation5MedianBin4
- Parallax_Error
- Detr_Skewness5_95



Matriz de Confusión

	Constant	Microvariable	Unsolved	Periodic		
Constant	212	549	1		Constant	
Microvariable	112	3738	94	405	11	Microvariable
Unsolved		319	313	152	10	Unsolved
Periodic		523	68	6167	196	Periodic
		33	27	503	2094	



Sample of Surveyed Stars

1. Variability detection

Variable Stars (10 to 30%)

2. Periodicity detection

Periodic Stars

Non-Periodic Stars

Constant Stars

ECL

LPV

CEP

RR

DSCT

SPB

ACV

3. Periodic Star Classification

I

Be

L

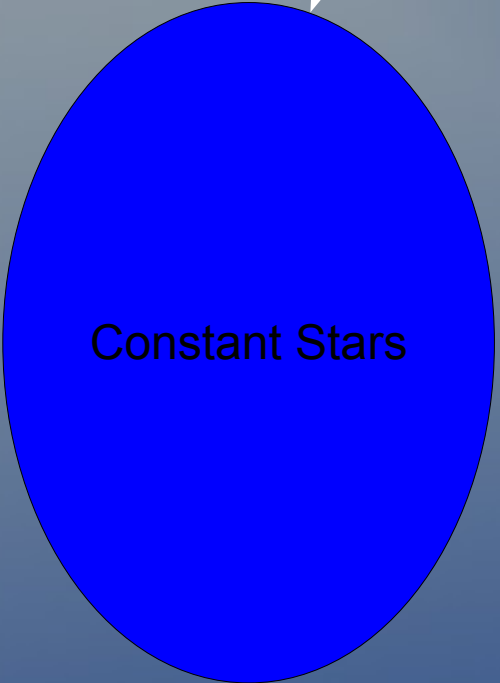
SR

RS

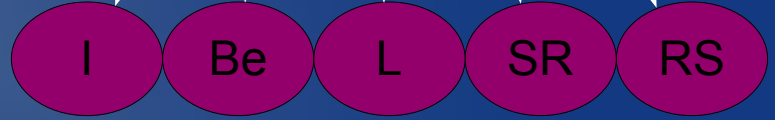
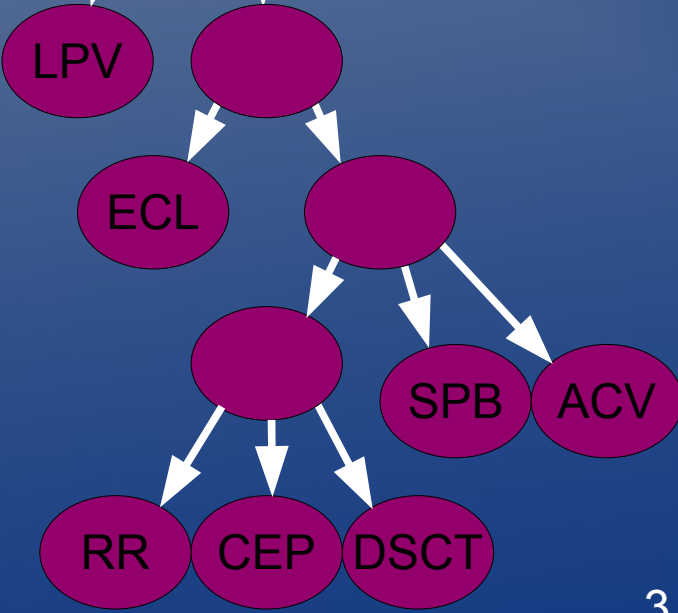
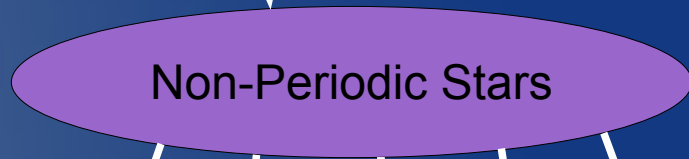
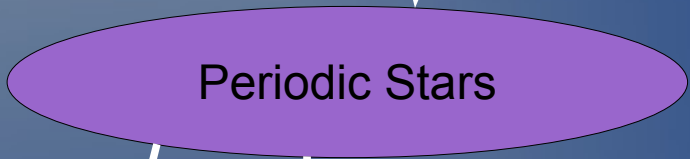
4. Non-Periodic Star Classification



1. Variability detection



2. Periodicity detection



4. Non-Periodic Star Classification

3. Periodic Star Classification

EA	EB	EW	ELL	LPV	RV	CWA	CWB	DCEP	DCEPS	CEP(B)	RRAB	RRC	GDOR	DSCT	DSCTC	BCEP	SPB	BE+GCAS	ACYG	ACV	SXARI	BY+RS	
214	13									1													EA
19	191	28	2	1				2					1		4		3		2	2			EB
	30	76							1														EW
	14			1									1		1		3			5		2	ELL
				285																			LPV
	1			1				2	1														RV
	2				1			5														1	CWA
	1						2	2	1														CWB
								183	5	1													DCEP
	1							11	17													2	DCEPS
	1							4		6													CEP(B)
	1										69	1					1						RRAB
	2	4									1	12		1									RRC
													27										GDOR
	1	1									1			32	12								DSCT
	1													1	77					2			DSCTC
	1	1													1	26	1						BCEP
			1													1	74		1	4			SPB
1									1								5		2	4			BE+GCAS
	1																	1	13	2		1	ACYG
	3								1				1				6			66			ACV
	2																2			3			SXARI
	1							1														33	BY+RS

Conclusiones

- Esquema completo para la clasificación de estrellas variables
- Optimizado para Hipparcos
- La clasificación en Hipparcos es relativamente fácil: ejemplos limpios y estrellas bien conocidas
- Esta clasificación podría ser completada con información adicional
 - Series temporales de velocidades radiales
- Sigüientes pasos: Aplicar este esquema a otros surveys.