IZSAM G.CAPORALE
TERAMO

# Studio sull'origine geografica e valutazione della contaminazione da metalli pesanti ed Arsenico nella mozzarella di bufala (IZS AM 06/12 RC)

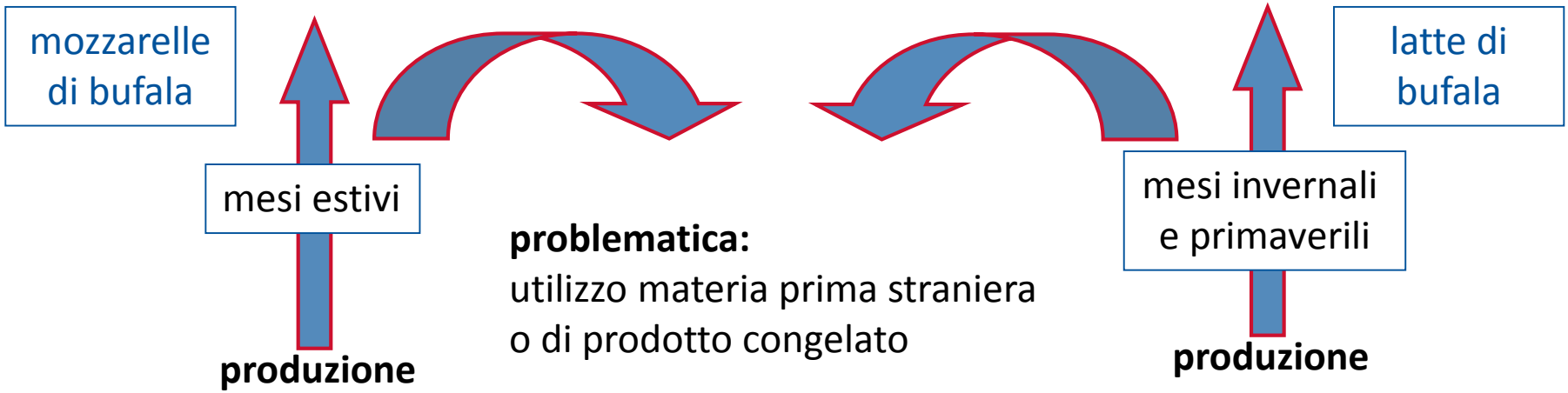
**Salvatore Milone**

Reparto Bromatologia, Residui alimenti per l'uomo e gli animali

# Introduzione

## Mozzarella di bufala

- mozzarella di latte di bufala ➲ 100% di latte di bufala
- mozzarella di bufala campana DOP ➲ 100% di latte di bufala proveniente da allevamenti posti all'interno della zona definita nel disciplinare di produzione
- mozzarella con latte di bufala ➲ latte bufalino e vaccino mescolati in varie percentuali




Camera dei Deputati, Atti Parlamentari, XVII Legislatura, Doc. XXII-*bis* N., 2015.
Relazione sulla contraffazione nel settore della Mozzarella di Bufala Campana

# Introduzione

Studi precedenti avevano mostrato buone performance di classificazione (predittività del modello del 99%) per origine geografica della mozzarella di bufala considerando:

- aziende costituite da allevamento e caseificio per la produzione di mozzarelle di bufala con latte proveniente esclusivamente dall'allevamento dell'azienda stessa.
- 10 caseifici in totale di cui 7 in un'area ristretta del casertano
- 162 campioni per la costruzione del modello di classificazione

**Situazione "reale"**

- non solo piccole realtà ma in un caseificio può confluire latte da differenti allevamenti posti in aree geografiche a volte molto diverse
- maggior numero di caseifici
- maggiore estensione territoriale
- maggior numero di campioni

# Obiettivi

❖ Studio basato su Spettroscopia di Risonanza Magnetica Nucleare in Alta Risoluzione del Protone ($^{1}H$ HR-NMR) in combinazione con analisi statistica multivariata per valutare la possibilità di classificare campioni di mozzarella di bufala per zona di produzione.

❖ Valutazione dei livelli di contaminazione da metalli pesanti (Piombo, Cadmio, Mercurio) ed Arsenico tramite ICP-MS (spettrometria di massa con plasma ad accoppiamento induttivo).

# Campionamento: considerazioni

☺ **Banca Dati Nazionale** pubblica relativa alla **presenza di capi bufalini** su territorio nazionale (BDN dell'Anagrafe Zootecnica istituita dal Ministero della Salute presso il CSN dell'Istituto "G. Caporale" di Teramo)

bufale in allevamenti bufalini

**74%** in Campania

**94%** nelle regioni interessate dalla DOP Mozzarella di Bufala Campana (al 30 aprile 2014)

☹ Elenco produttori di mozzarelle di bufala

☹ Tracciabilità latte bufalino

☺ Elenco caseifici iscritti al Consorzio di tutela della DOP "Mozzarella di bufala Campana"

riferimento per i campionamenti

☹ Mercato dinamico

- cessazione attività caseifici
- caseifici entrano ed escono dalla DOP

# Considerazioni

☹ Elenco produttori di mozzarelle di bufala

☹ Tracciabilità latte bufalino

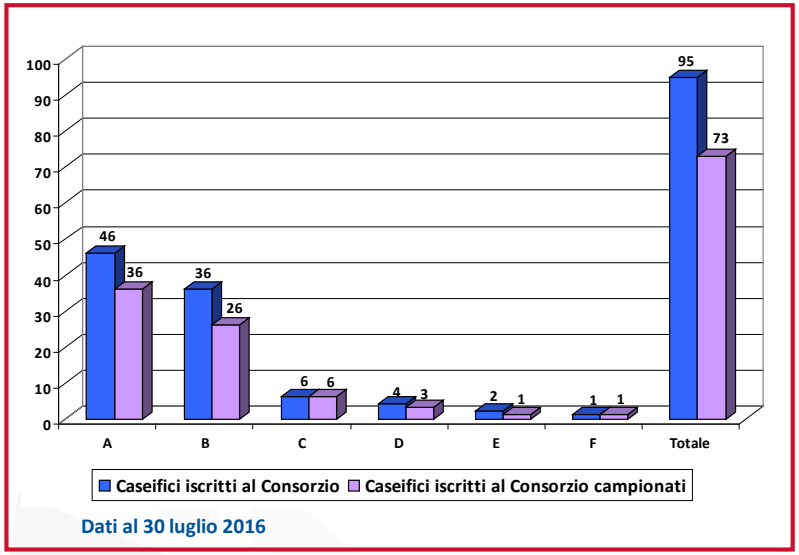
Decreto del 9 settembre 2014
Gazzetta Ufficiale Anno 155° Numero 219

È istituita la piattaforma informatica
"**Tracciabilità della filiera bufalina**"

☺

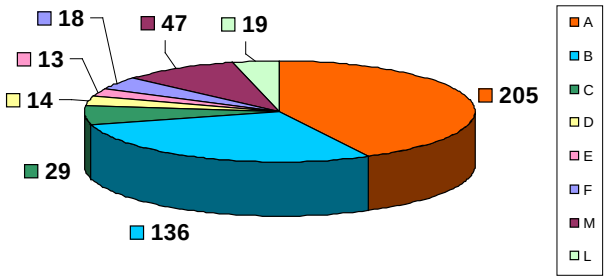
# Campionamento

**481 campioni**
prelevati in **115 caseifici o punti vendita**
periodo maggio 2014-maggio 2016

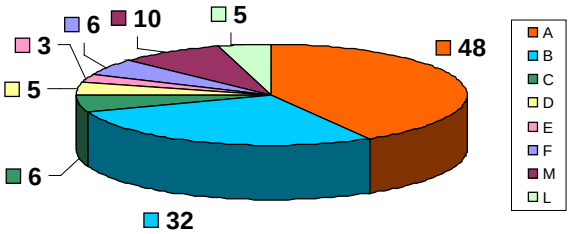



Dati al 30 luglio 2016

Campioni di mozzarella di bufala prelevati per zona di produzione




Numero di caseifici interessati dal campionamento



# Livelli di contaminazione per Piombo, Cadmio, Mercurio e Arsenico

## Metodi ufficiali utilizzati:

- ✓ preparazione del campione: UNI EN 13804:2013
- ✓ mineralizzazione del campione: UNI EN 13805:2014
- ✓ analisi in ICP-MS (spettrometria di massa con plasma ad accoppiamento induttivo): UNI EN 15763:2010

| | Piombo | Cadmio | Mercurio | Arsenico |
|---|---|---|---|---|
| **LOD** (mg/kg) | 0.010 | 0.0010 | 0.0050 | 0.010 |
| **LOQ** (mg/kg) | 0.020 | 0.0020 | 0.010 | 0.020 |

•LOD: limite di rivelazione
•LOQ: limite di quantificazione

# Livelli di contaminazione per Piombo, Cadmio, Mercurio e Arsenico

**Non è stato stabilito un tenore massimo accettabile da alcuna legislazione europea o nazionale**

**Valutazione**

in termini di possibile impatto sulla salute pubblica in base alla dose ammissibile giornaliera

# Livelli di contaminazione per Piombo, Cadmio, Mercurio e Arsenico

Analisi effettuate su 198 mozzarelle di bufala

Riferimenti per calcolo di quantità di contaminante assunto giornalmente:

ricerca dell'Istituto Nazionale di Ricerca per l'Alimentazione e la Nutrizione (INRAN) per peso corporeo e consumo di mozzarella di bufala

età (10-18 anni), peso corporeo medio (52,6 kg). Quantità consumata giornalmente dell'alimento al 95° percentile pari a 41,7 g/die

**non è mai superata la dose ammissibile giornaliera**

# Analisi $^{1}H$ HR-NMR

omogeneizzazione

2 g
omogeneizzato

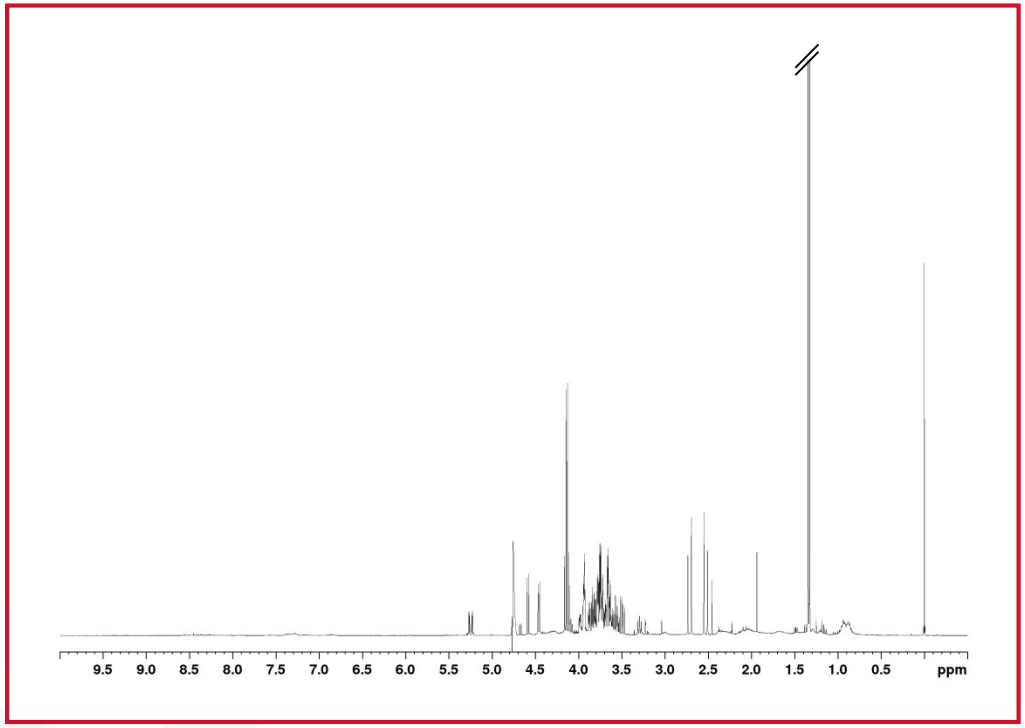
estrazione
in $D_2O$

centrifugazione e
filtrazione (0.2 µm)

estratto filtrato
+ $NaN_3$+TSP

**NMR Bruker**
**AVANCE III 400 Ultrashield**

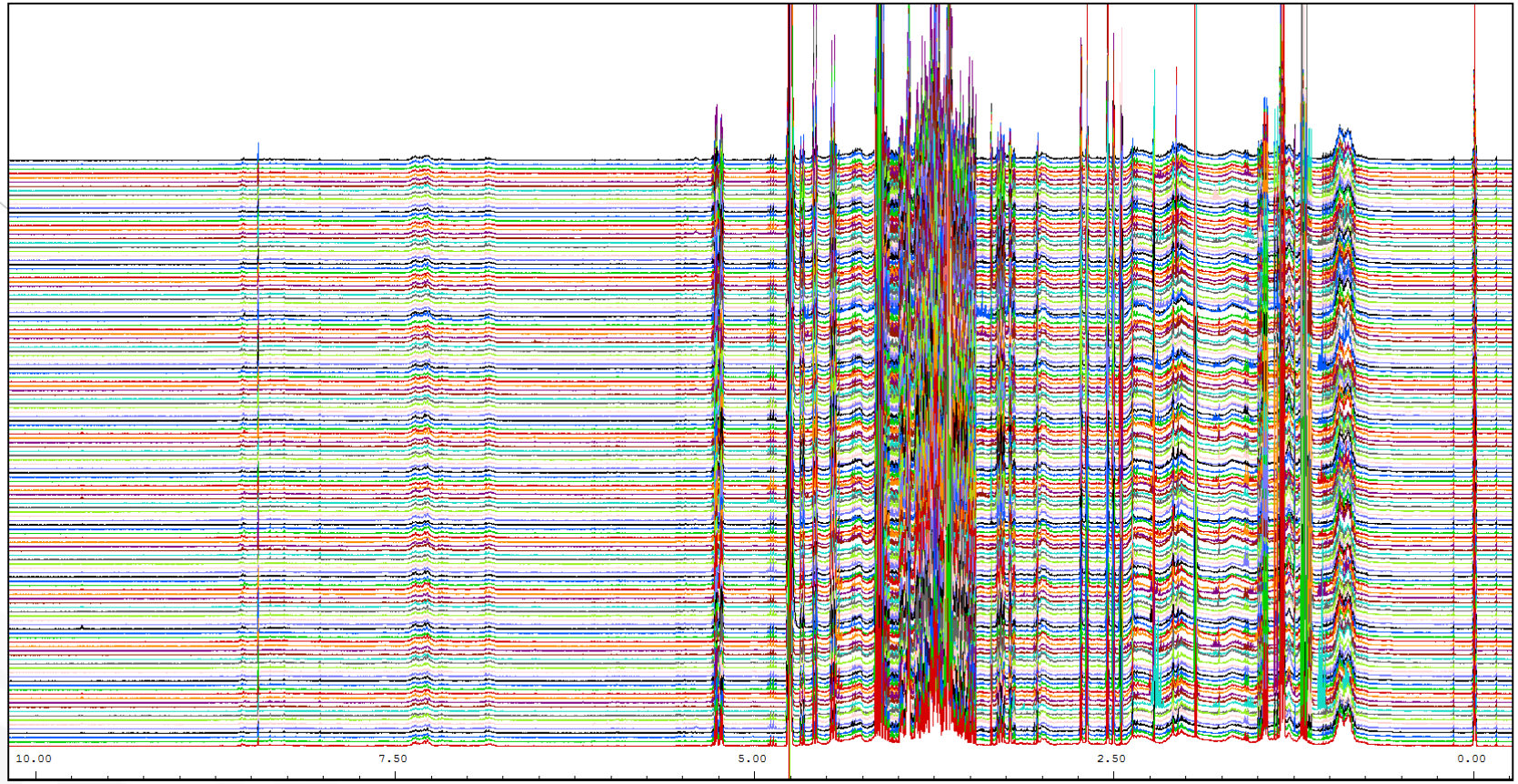
# Analisi $^{1}H$ HR-NMR




esempio di spettro $^{1}H$ HR-NMR
di un estratto di mozzarella di bufala in $D_2O$

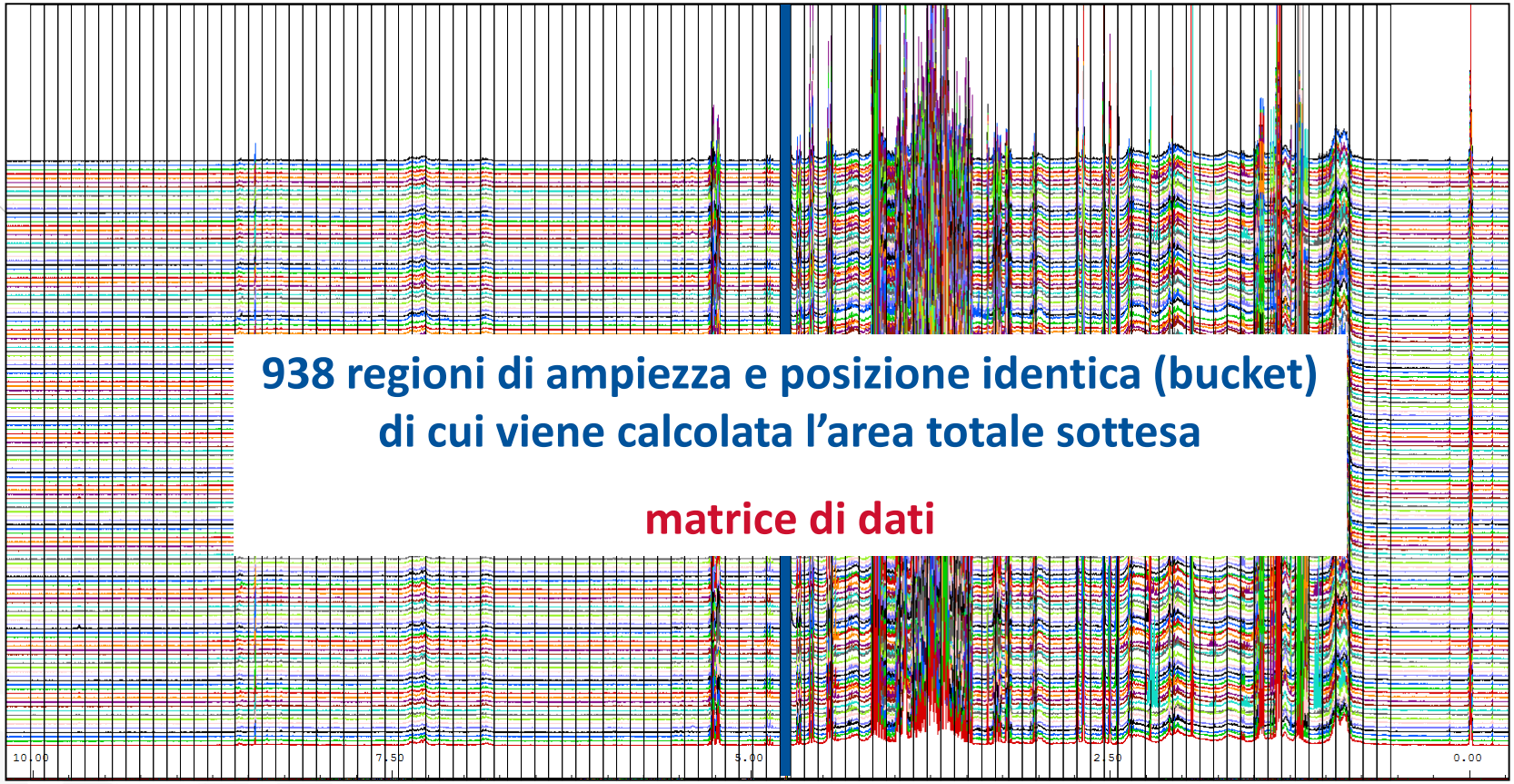
**NMR Bruker**
**AVANCE III 400 Ultrashield**

# Spettri $^{1}H$ HR-NMR e pre-processing




Sovrapposizione di spettri $^{1}H$ HR-NMR di estratti di mozzarella di bufala in $D_2O$

# Spettri $^{1}$H HR-NMR e pre-processing



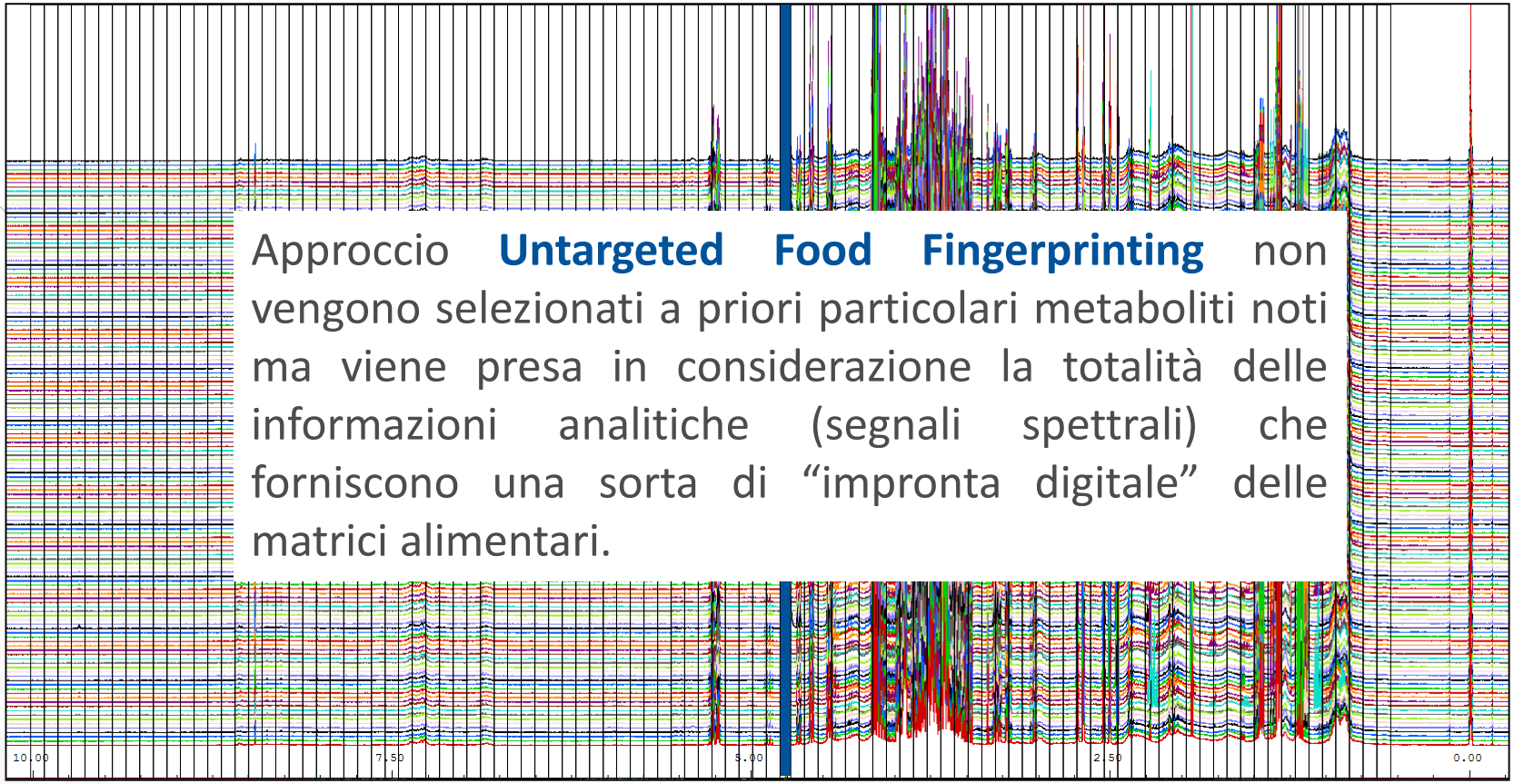

**938 regioni di ampiezza e posizione identica (bucket) di cui viene calcolata l'area totale sottesa**

**matrice di dati**

**Bucketing rettangolare**

Bucket: ampiezza 0.01 ppm, zona spettrale (10-0.5) ppm escluso segnale residuo dell'acqua

# Approccio




Approccio **Untargeted Food Fingerprinting** non vengono selezionati a priori particolari metaboliti noti ma viene presa in considerazione la totalità delle informazioni analitiche (segnali spettrali) che forniscono una sorta di “impronta digitale” delle matrici alimentari.

**Bucketing rettangolare**

# Analisi statistica multivariata applicata a dati NMR

selezione delle variabili Random Forest → Linear Discriminant Analysis (LDA) → Modelli di classificazione

**La LDA è una tecnica di analisi statistica multivariata supervised.**

**Le tecniche supervised si basano sull'ipotesi a priori che esistano dei gruppi all'interno dei campioni analizzati.**

- **380 campioni**
  - costruzione modelli
  - validazione interna (Cross Validation-Leave One Out, CV-LOO)

➲ **calcolo performance di classificazione e predizione dei campioni del modello**

- **101 campioni**

➲ **calcolo performance di predizione dei campioni esterni sul modello ottenuto**

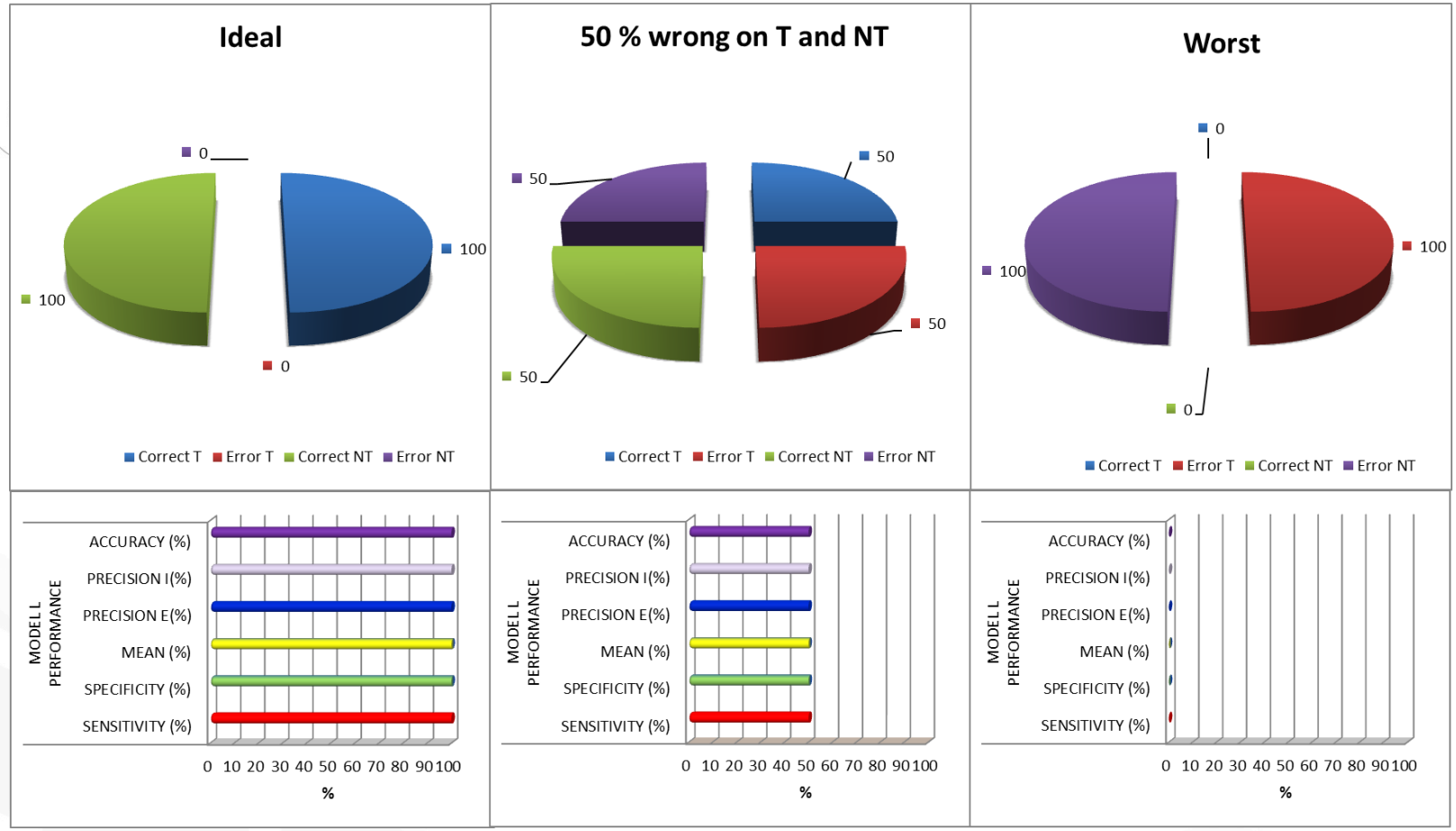
Software R (R version 3.1.0 (2014-04-10) "Spring Dance" Copyright (C) 2014 The R Foundation for Statistical Computing Vienna, Austria)

# Come si legge

## “T” Target - ”NT” None Target

### Matrici di confusione della LDA e performance del modello ottenuto

**Matrici simmetriche a due classi: I tre casi per capire**

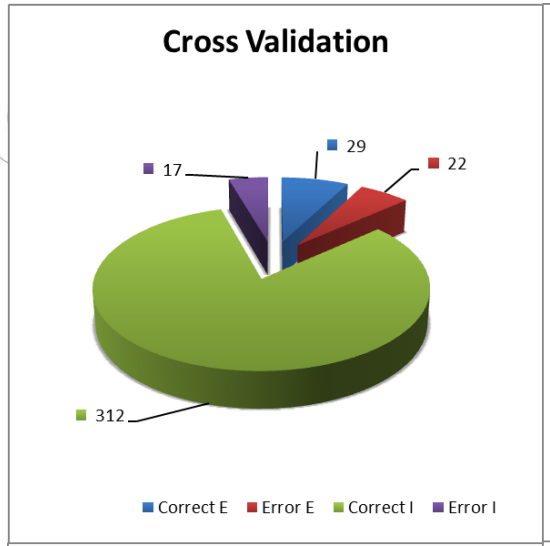


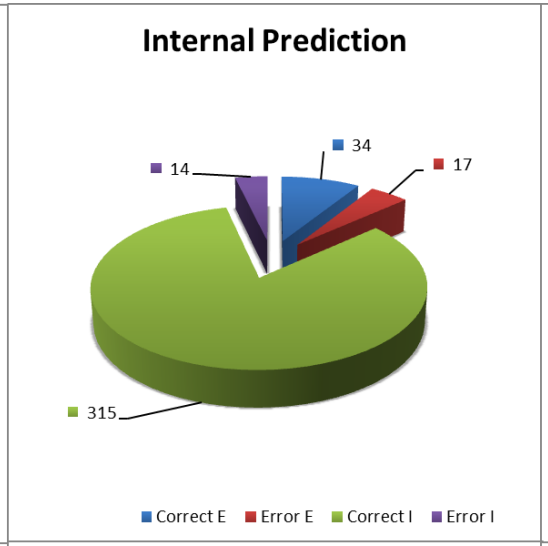
# MODELLO EI

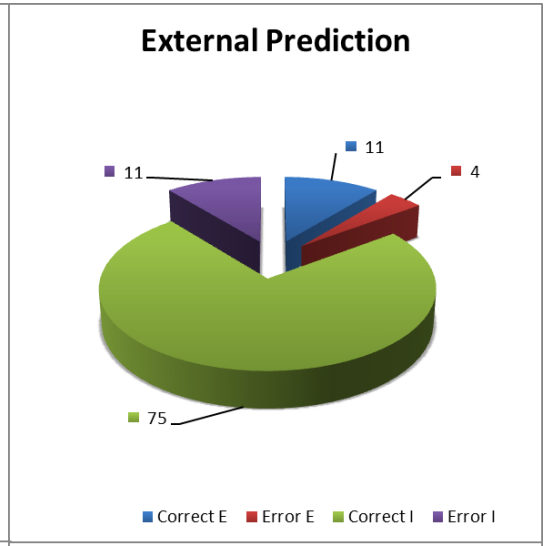
## “Esterni area DOP”-”Interni area DOP”

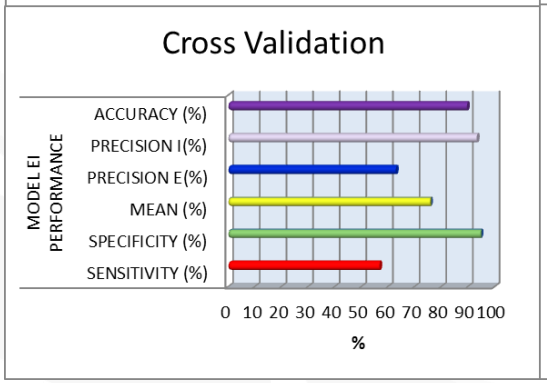
### Matrici di confusione della LDA e performance del modello ottenuto

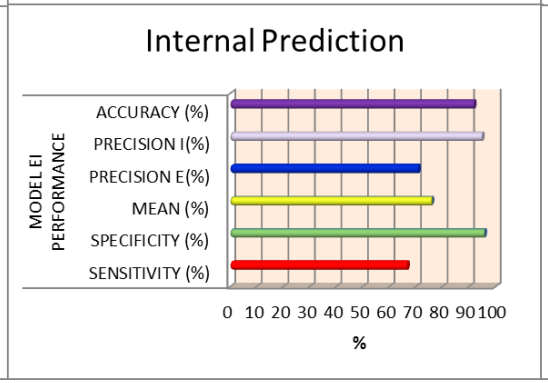
**13 variabili selezionate tramite Random Forest**

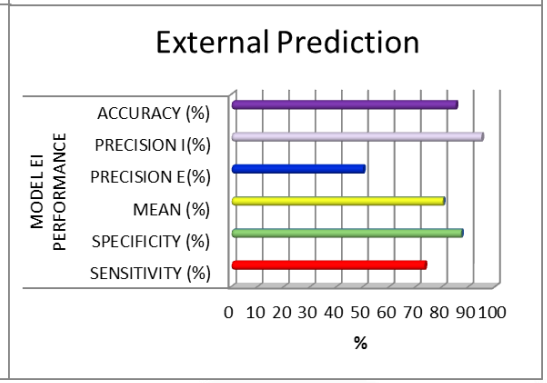


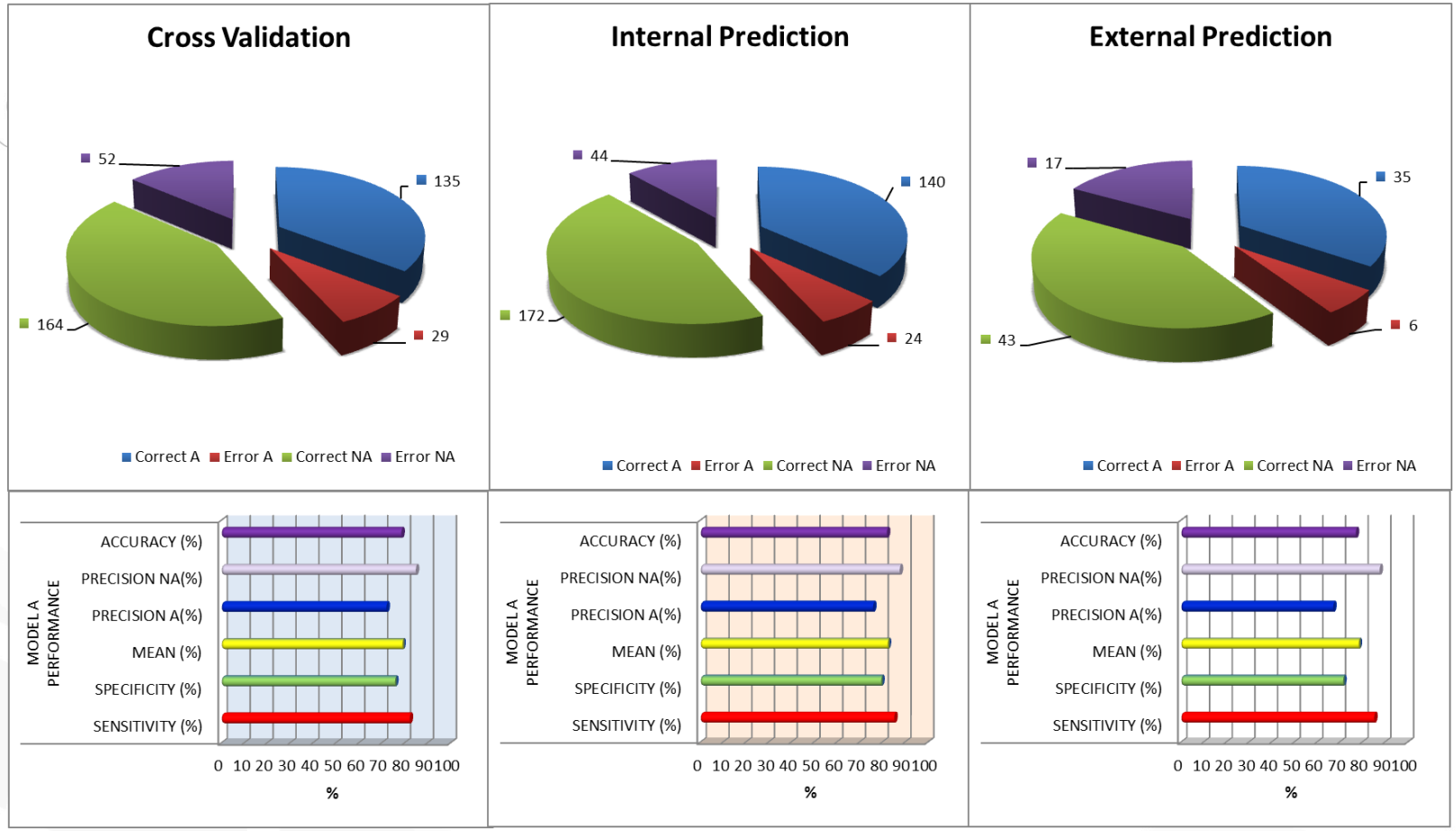



# MODELLO A

## “Zona A”-”Zona NA”

### Matrici di confusione della LDA e performance del modello ottenuto

**26 variabili selezionate tramite Random Forest**

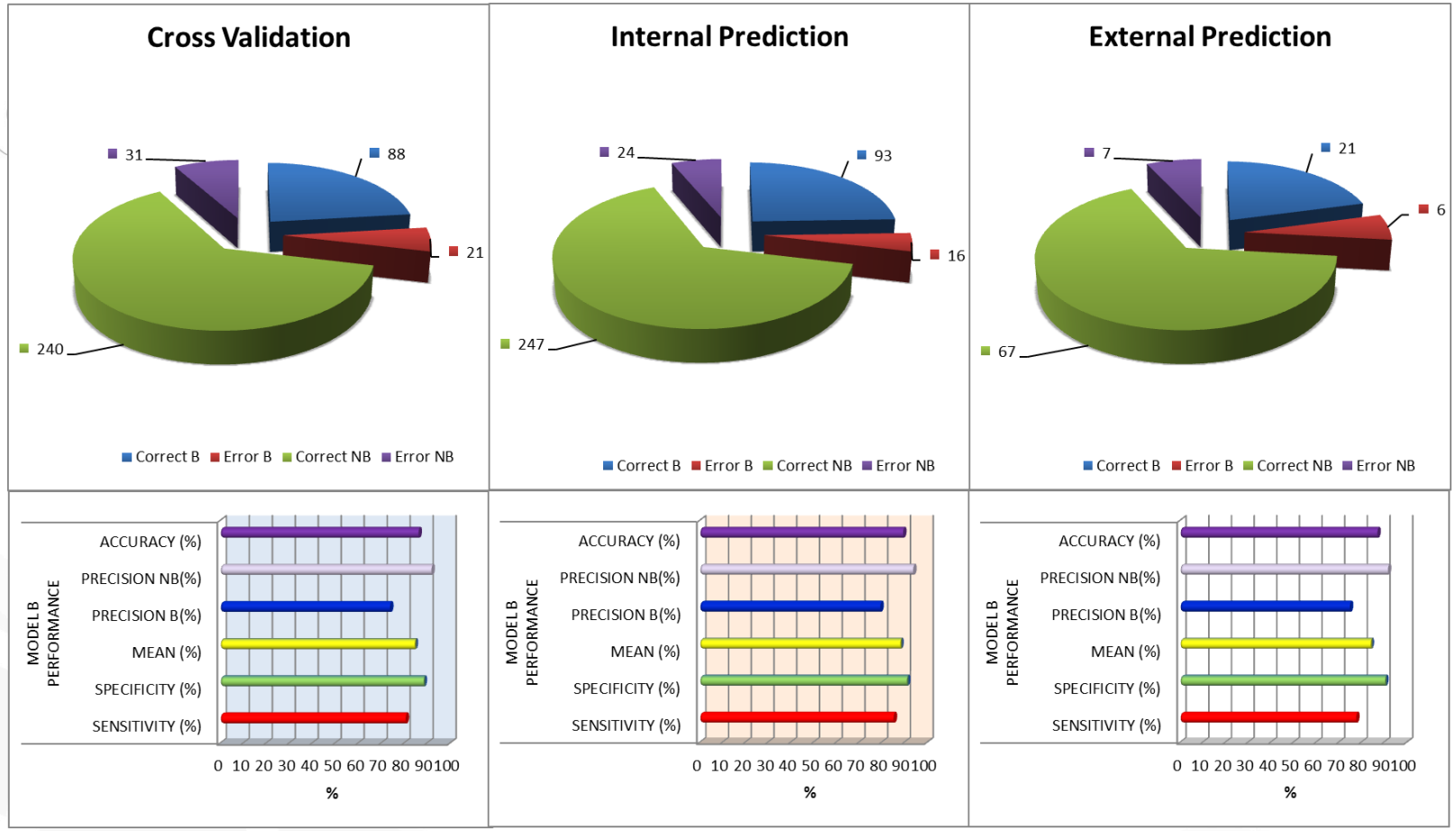


# MODELLO B

## “Zona B”-”Zona NB”

### Matrici di confusione della LDA e performance del modello ottenuto

**16 variabili selezionate tramite Random Forest**

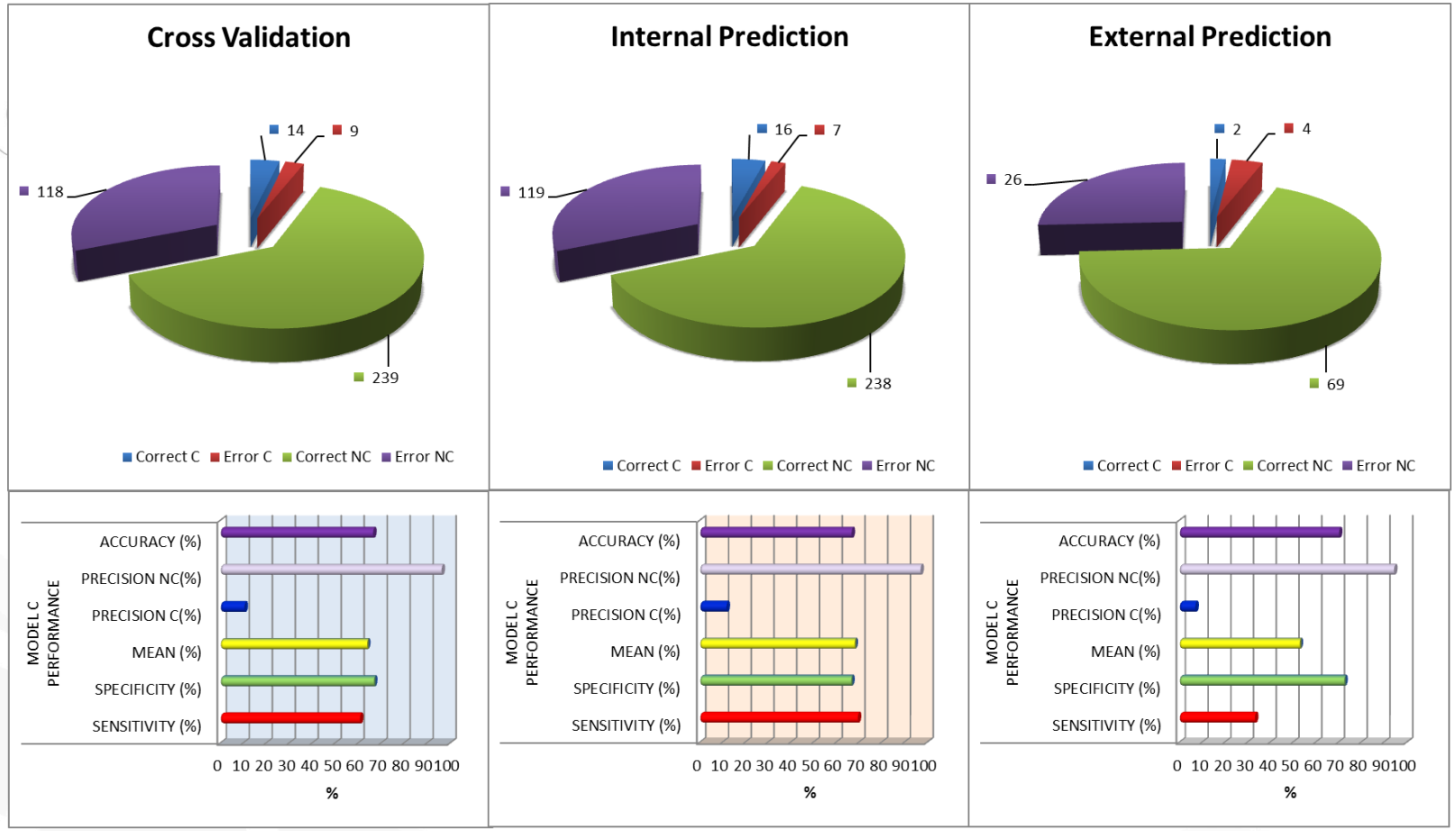


# MODELLO C

## “Zona C”-”Zona NC”

### Matrici di confusione della LDA e performance del modello ottenuto

**3 variabili selezionate tramite Random Forest**

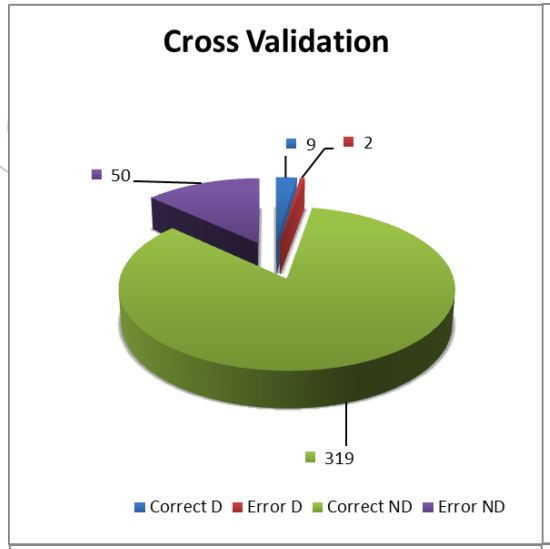


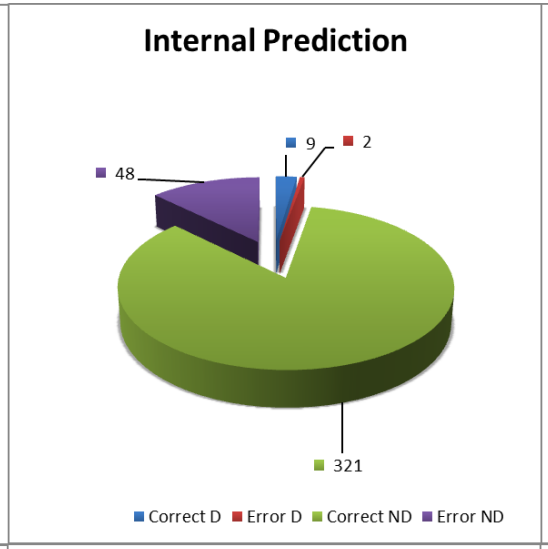
# MODELLO D

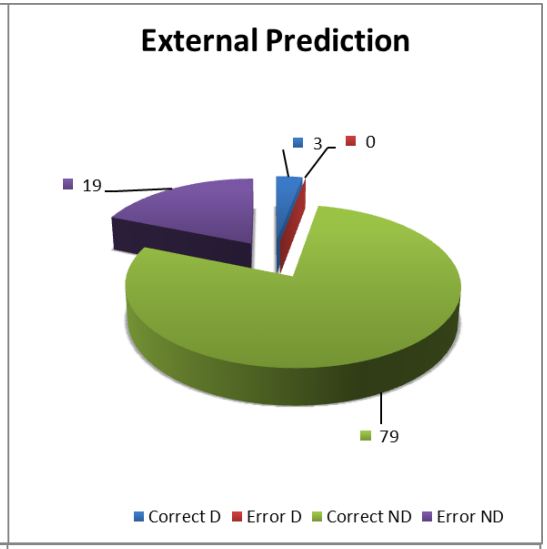
## “Zona D”-”Zona ND”

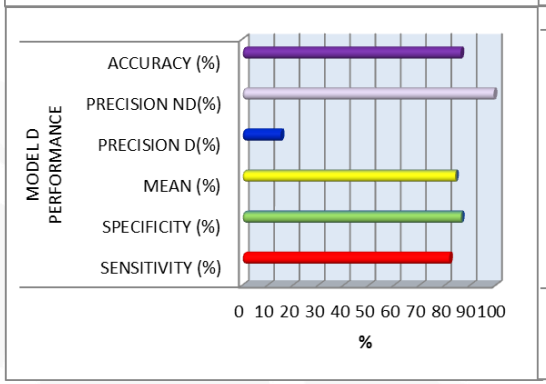
### Matrici di confusione della LDA e performance del modello ottenuto

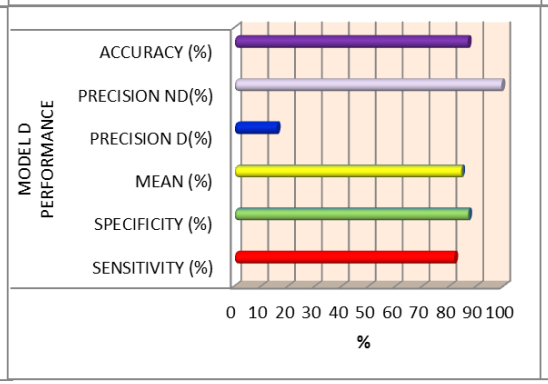
**6 variabili selezionate tramite Random Forest**

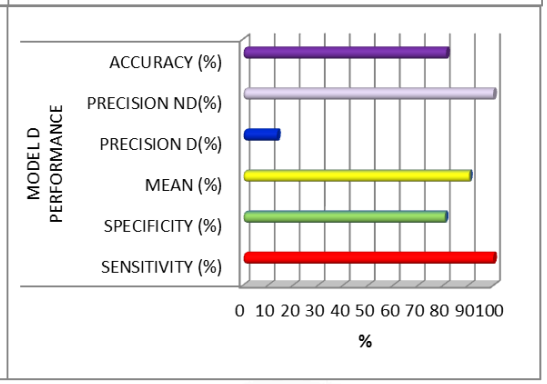


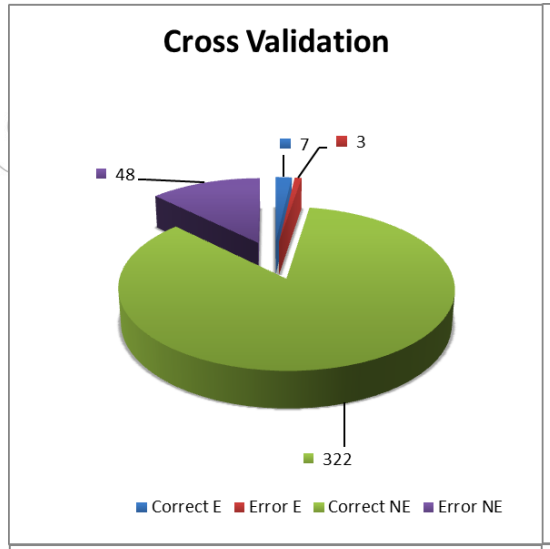



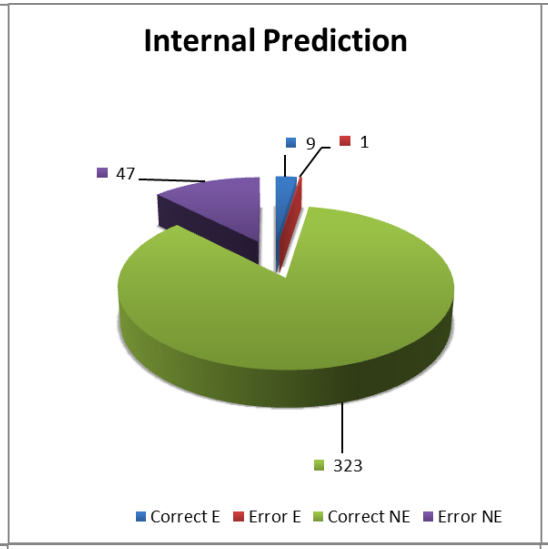
# MODELLO E

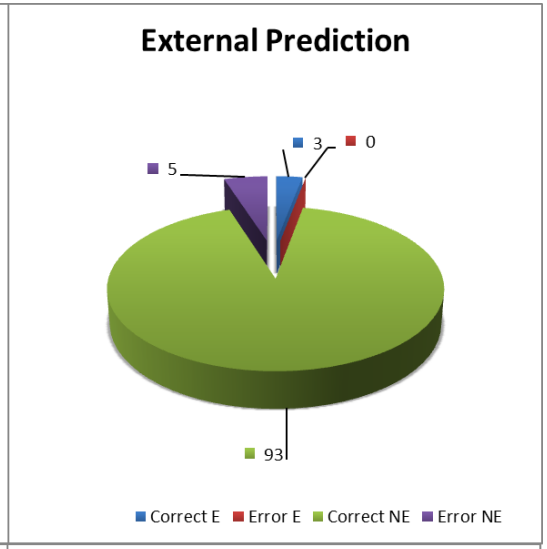
## “Zona E”-”Zona NE”

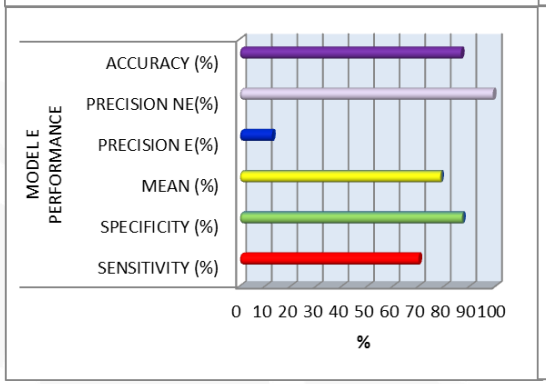
### Matrici di confusione della LDA e performance del modello ottenuto

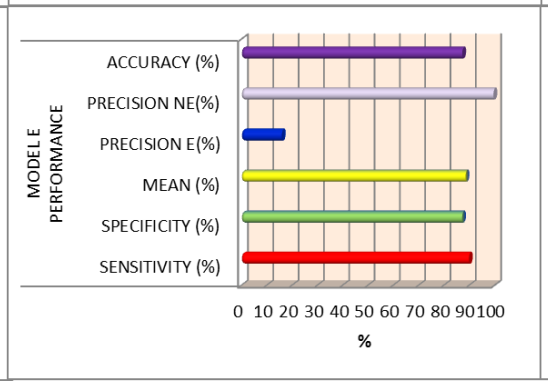
**4 variabili selezionate tramite Random Forest**

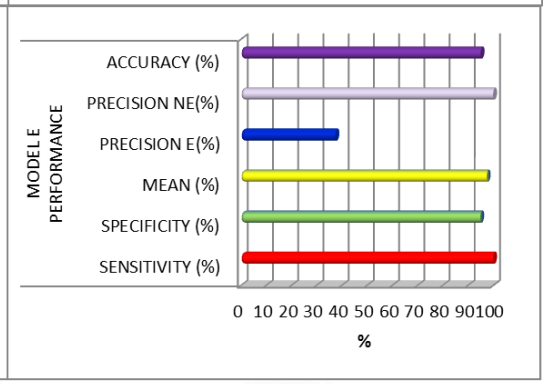


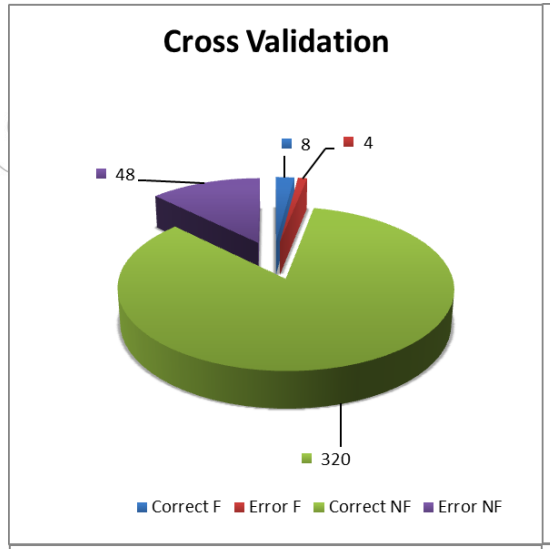



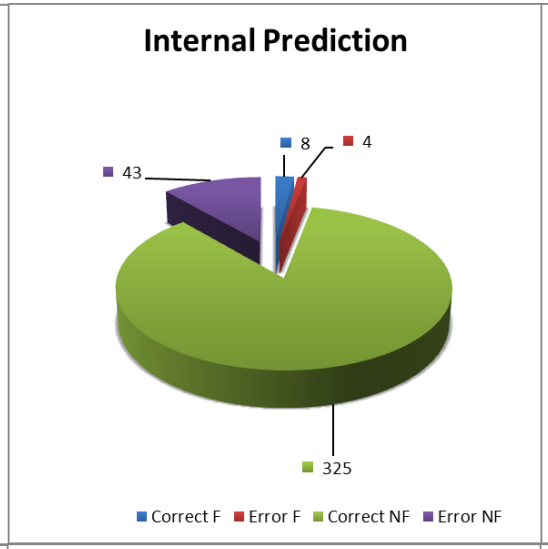
# MODELLO F

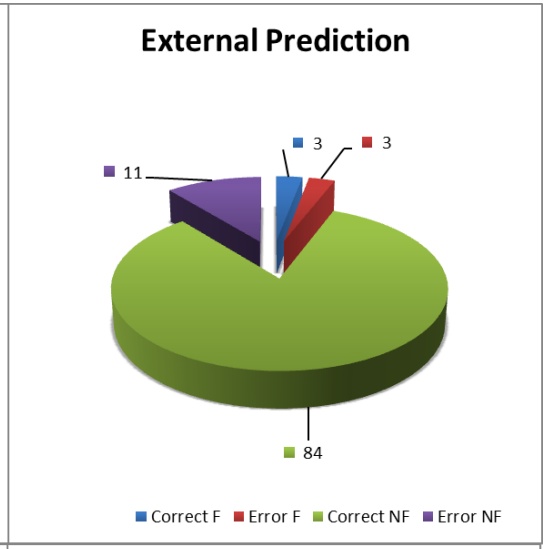
## “Zona F”-”Zona NF”

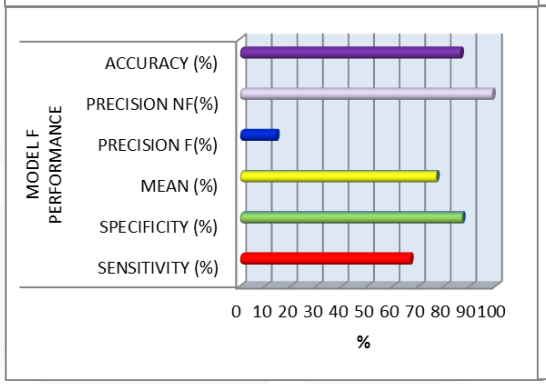
### Matrici di confusione della LDA e performance del modello ottenuto

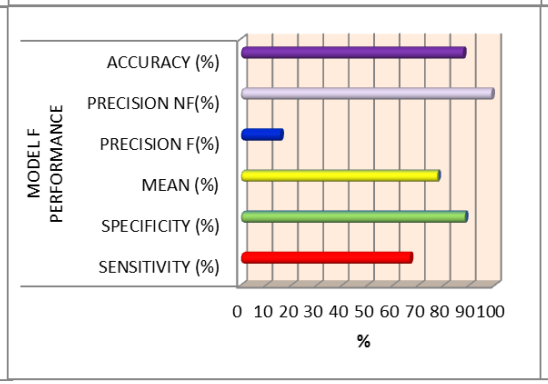
**3 variabili selezionate tramite Random Forest**

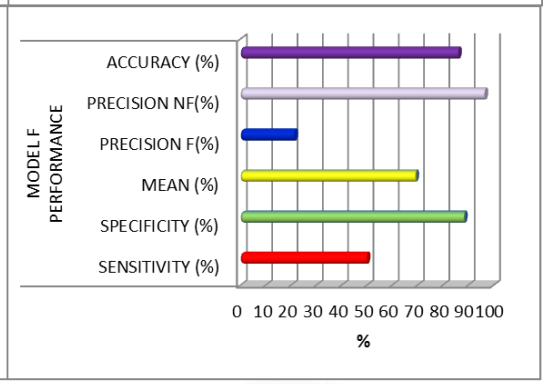


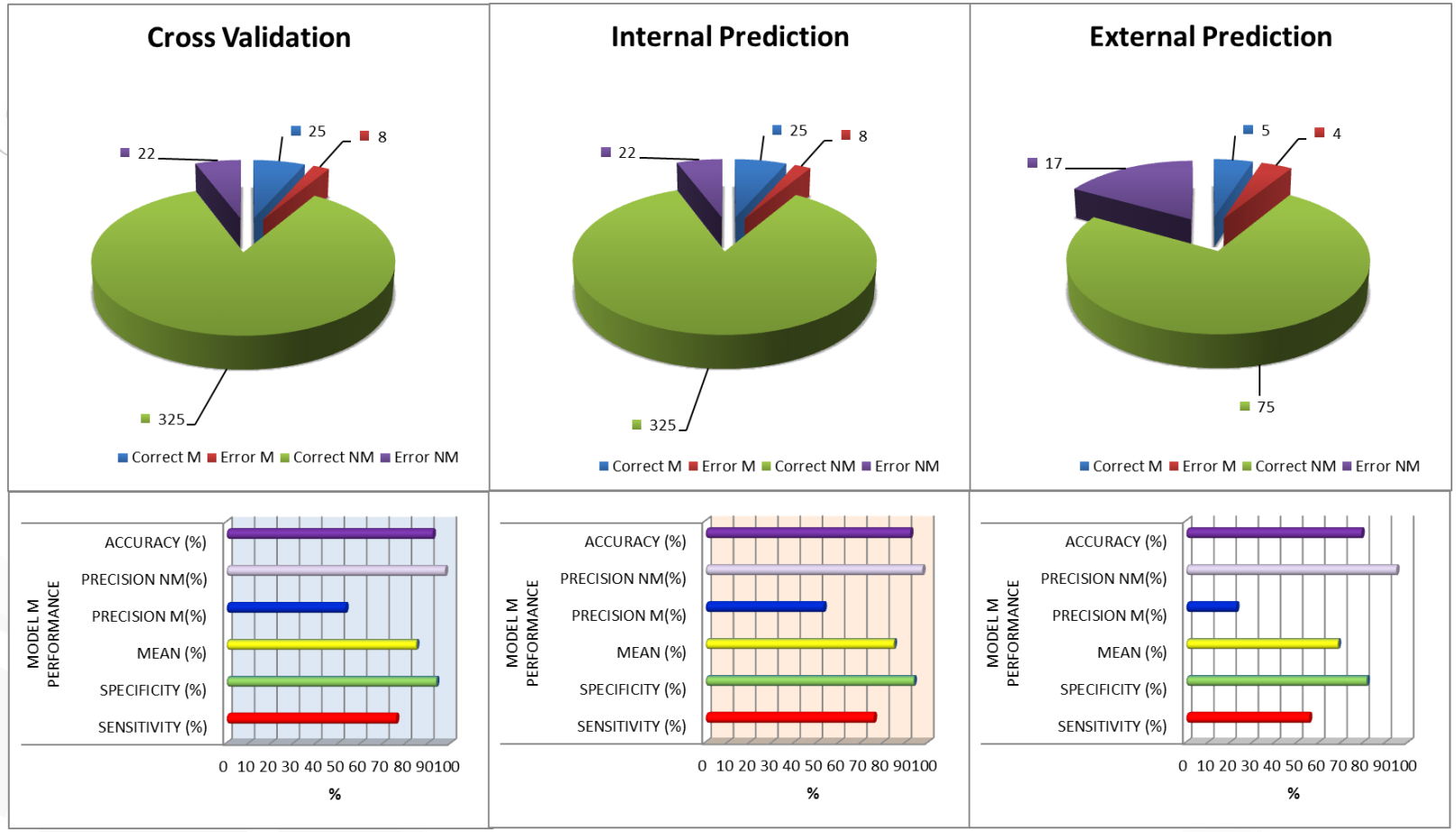



# MODELLO M

## “Zona M”-”Zona NM”

### Matrici di confusione della LDA e performance del modello ottenuto

**4 variabili selezionate tramite Random Forest**

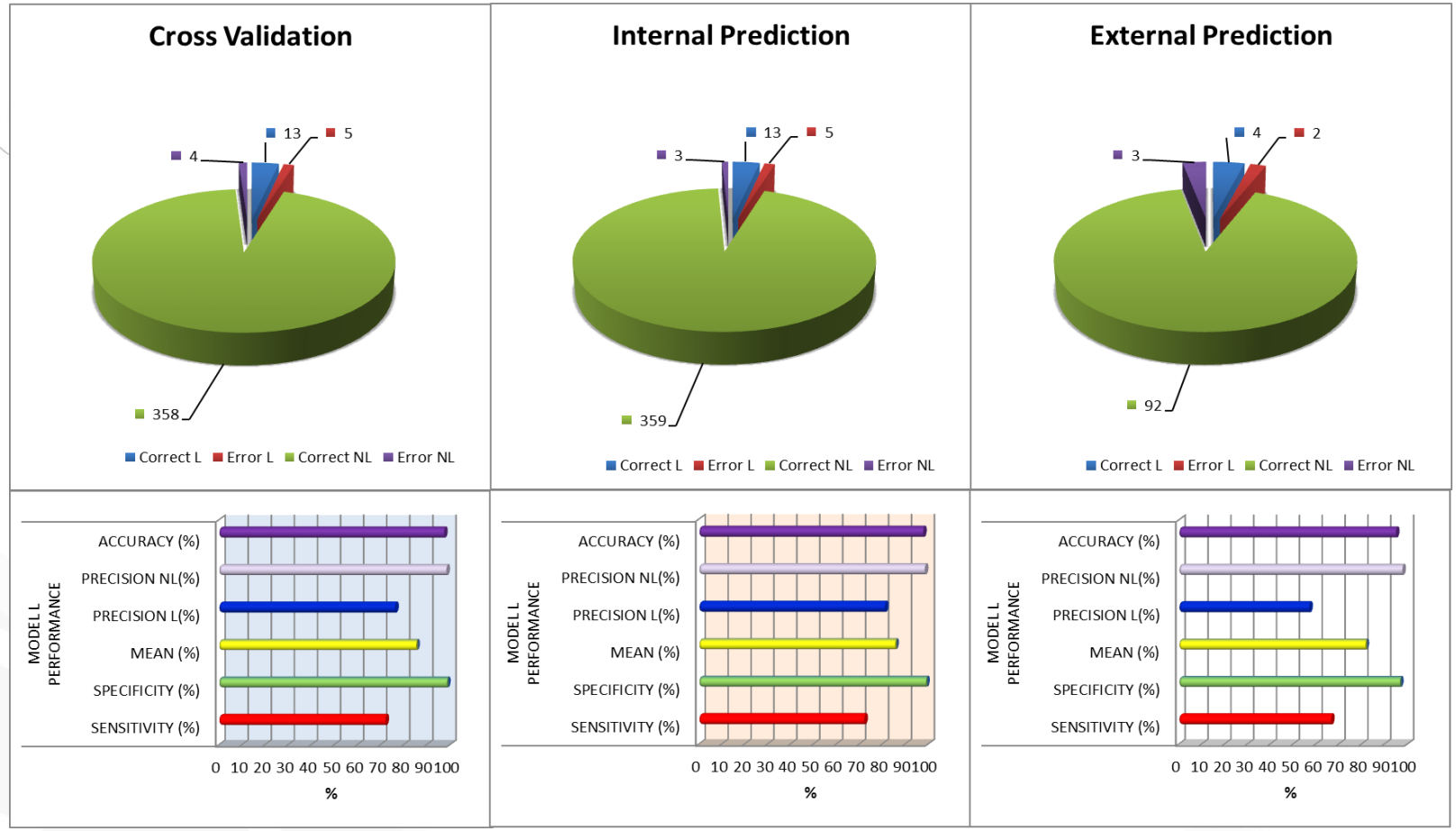


# MODELLO L

## “Zona L”-”Zona NL”

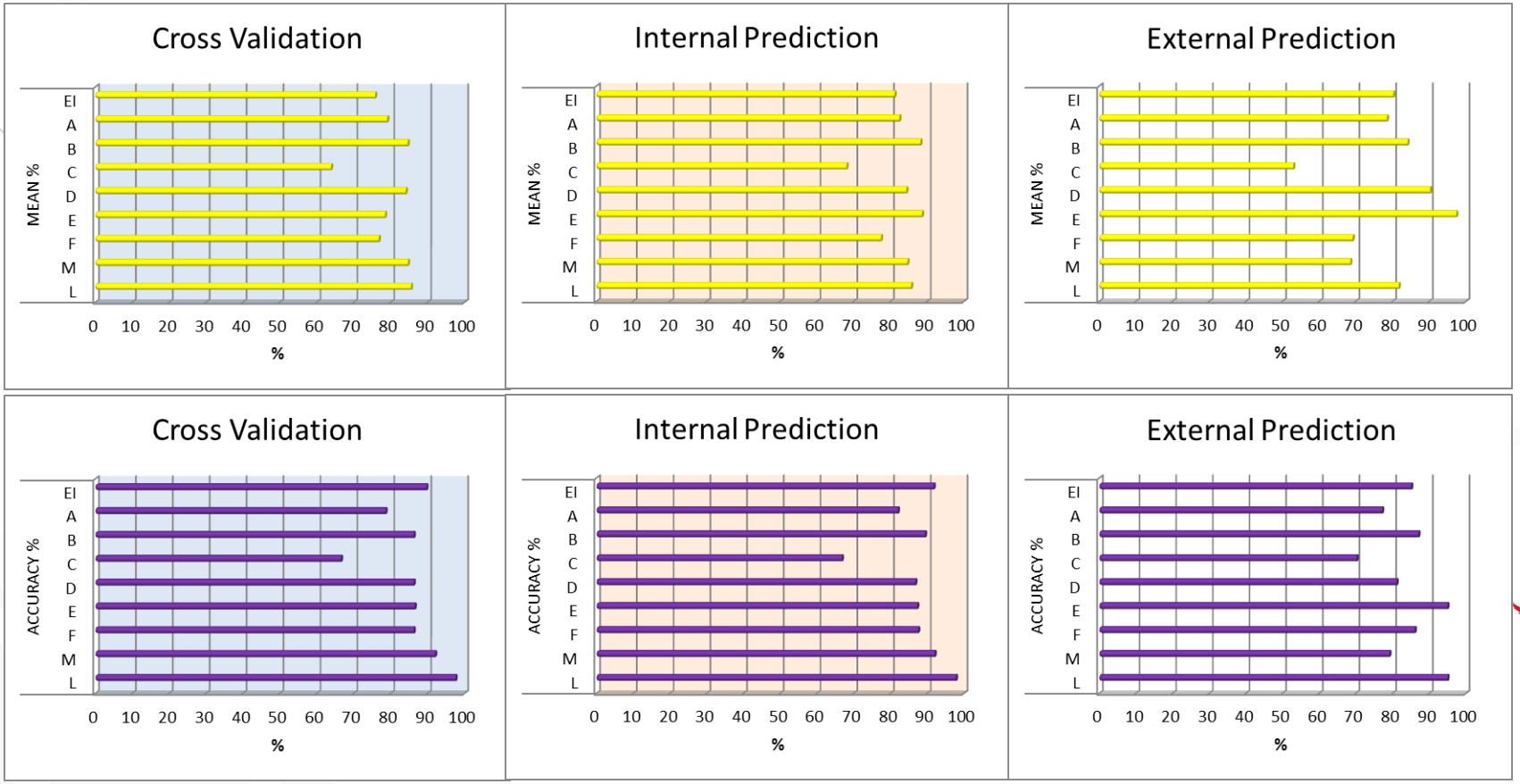
### Matrici di confusione della LDA e performance del modello ottenuto

**5 variabili selezionate tramite Random Forest**



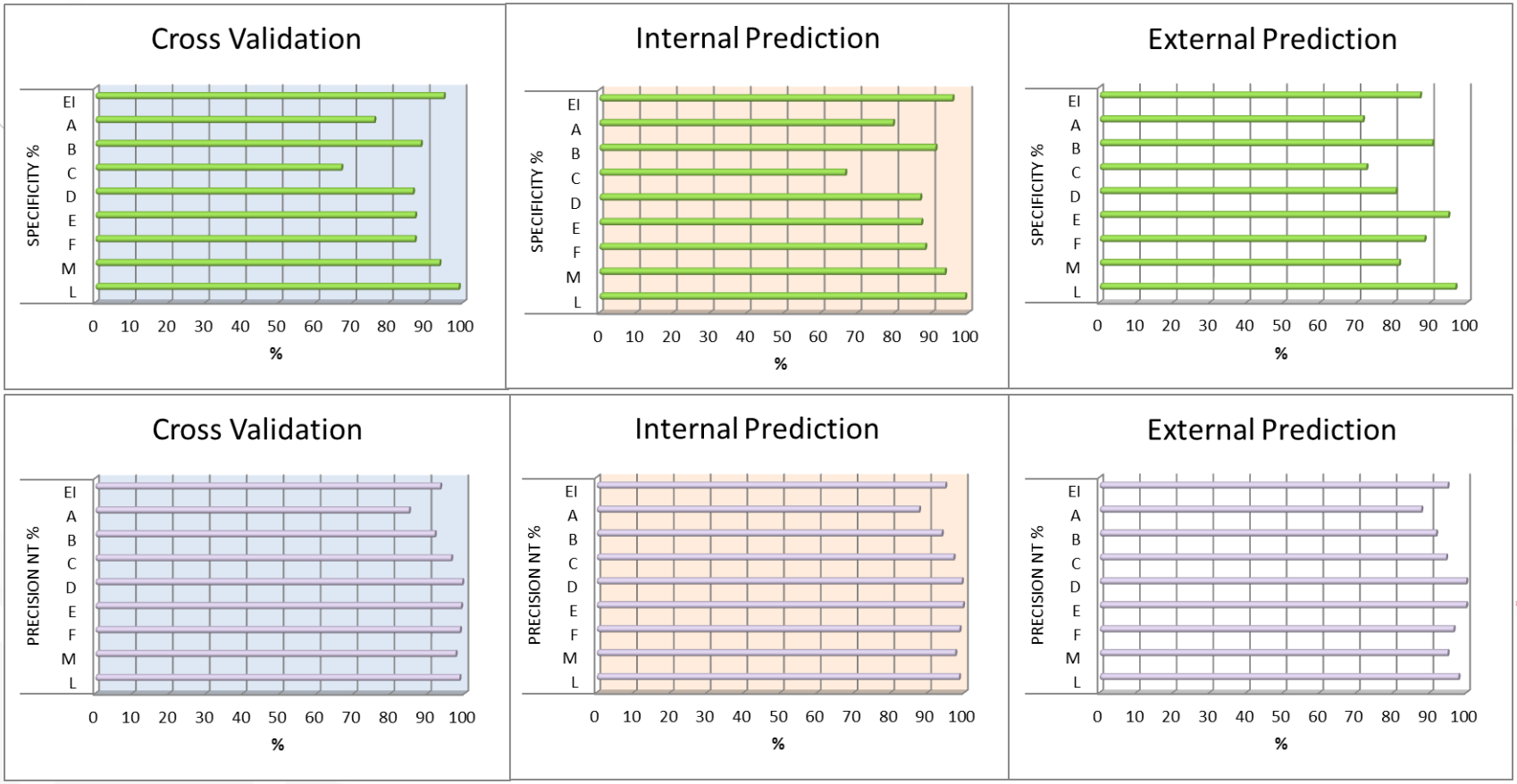
# RIASSUMENDO

## CONFRONTI PERFORMANCE DEI MODELLI



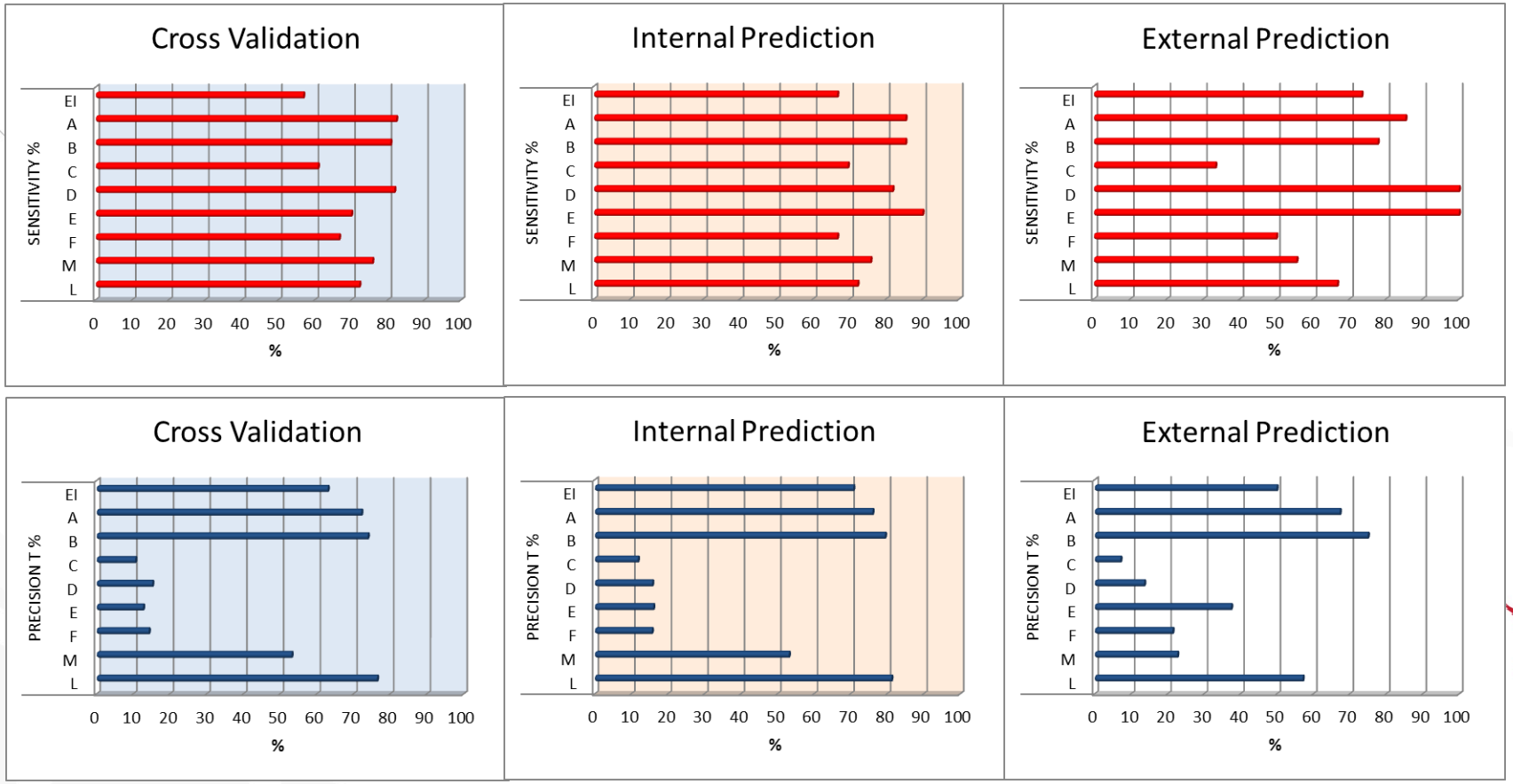
# RIASSUMENDO

## CONFRONTI PERFORMANCE DEI MODELLI (NT)



# RIASSUMENDO

## CONFRONTI PERFORMANCE DEI MODELLI (T)

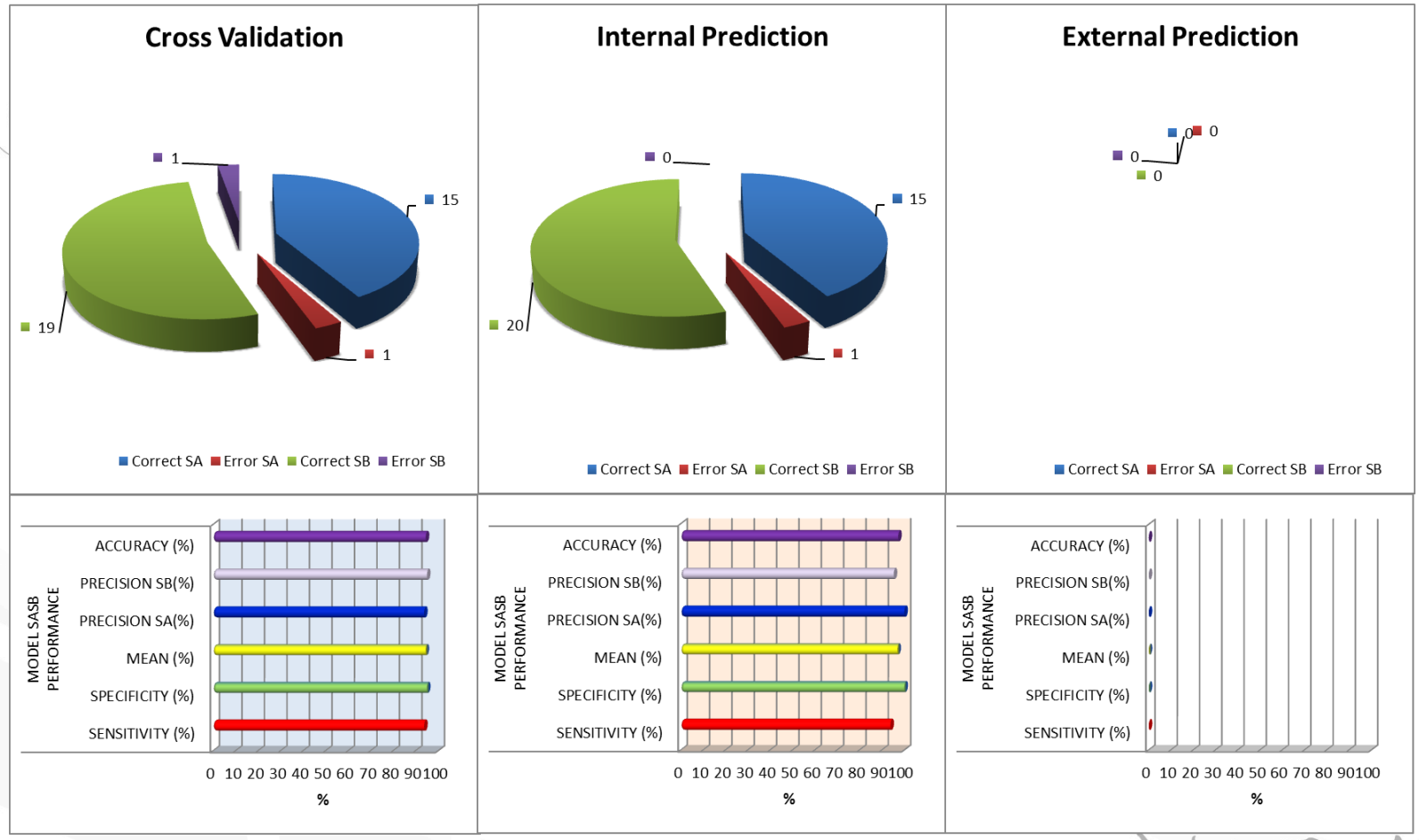


# MODELLO SASB

## “Zona SA”-”Zona SB”

### Matrici di confusione della LDA e performance del modello ottenuto

**7 variabili selezionate tramite Random Forest**

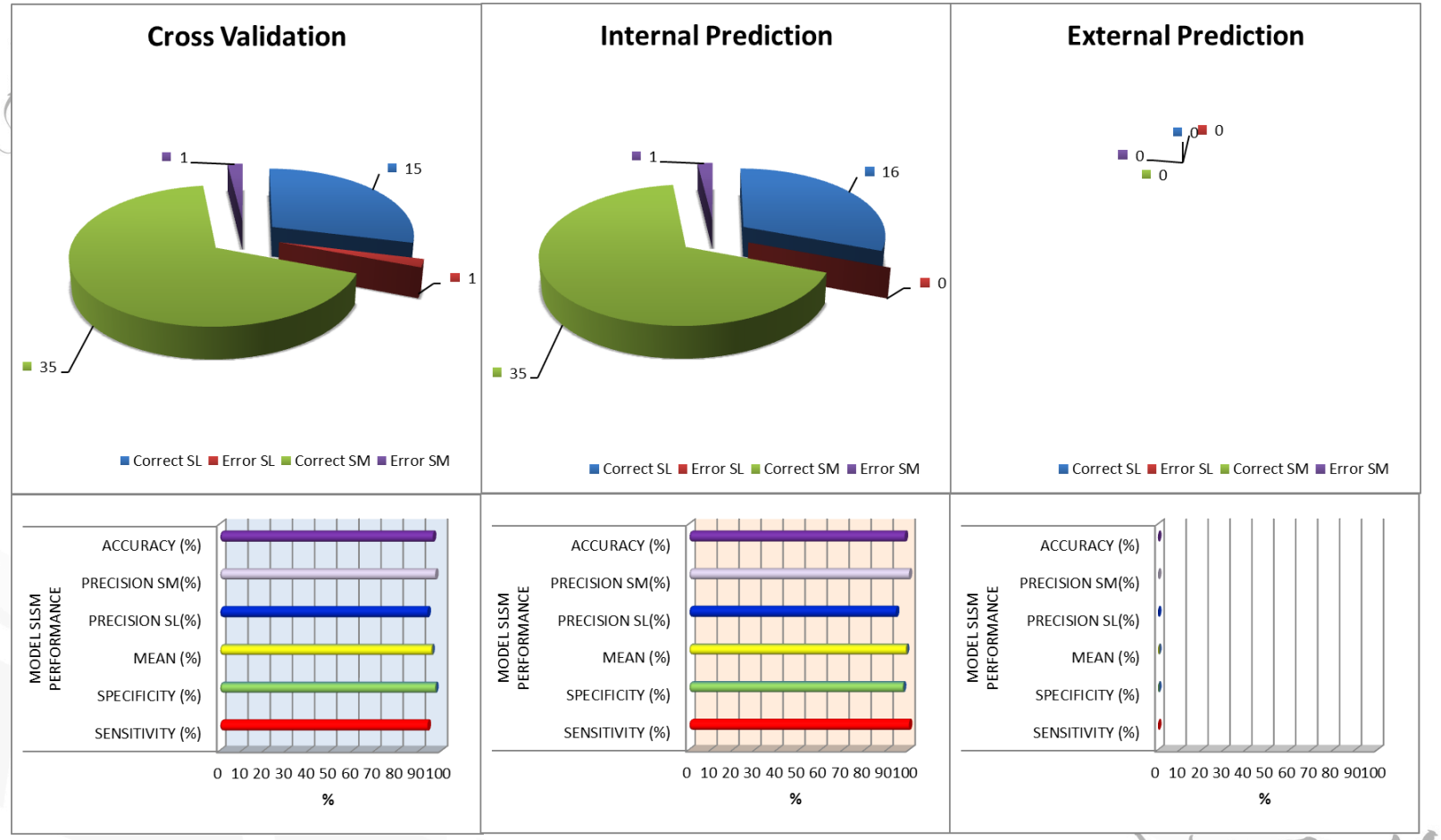


# MODELLO SLSM

## “Zona SL”-”Zona SM”

### Matrici di confusione della LDA e performance del modello ottenuto

**12 variabili selezionate tramite Random Forest**



# Conclusioni e sviluppi futuri

❖ Studio basato su spettroscopia di Risonanza Magnetica Nucleare ad Alta Risoluzione del Protone ($^{1}H$ HR-NMR) in combinazione con analisi statistica multivariata per valutare la possibilità di classificare campioni di mozzarella di bufala per zona di produzione.

I risultati ottenuti per un sistema così complesso sono molto interessanti ed incoraggianti.
Le indagini NMR effettuate hanno consentito di avviare la creazione di una banca dati di profili spettroscopici della matrice in esame e di valutare tramite modelli di classificazione la possibilità di discriminare i campioni sulla base dell'origine geografica.
Si sono ottenute molte informazioni per poter migliorare notevolmente il disegno sperimentale di prossime ricerche.
Una continua implementazione della banca dati con nuovi campioni e dati relativi alla tracciabilità ottenuti lavorando in sinergia con enti preposti e la combinazione di dati analitici provenienti da diverse tecniche potrebbero permettere lo sviluppo di nuovi modelli predittivi sempre più potenti.

# ELENCO DEI COLLABORATORI

- Rocco Mauro Ciccone (reperimento campioni)
- Antonio D'Aloise (reperimento campioni)
- Francesca Marchegiani (valutazione  contaminazione metalli pesanti e As)
- Valeria Melai (valutazione  contaminazione metalli pesanti e As)
- Alessandra Micozzi (studio origine geografica tramite $^{1}$H HR-NMR)

# Grazie per l’attenzione

**Contatti:**

**s.milone@izs.it**