

Χωρο-χρονικά μοντέλα για την μελέτη της επιδημίας της ευλογιάς του πρόβατου

Μαλέσιος Χρυσοβαλάντης*, Δεμίρης Νικόλαος² και Κουτρομανιδής Θεόδωρος¹

Τμήμα Αγροτικής Ανάπτυξης, Δημοκρίτειο Πανεπιστήμιο Θράκης

Τμήμα Στατιστικής, Οικονομικό Πανεπιστήμιο Αθηνών

Περίληψη

Στην παρούσα μελέτη χρησιμοποιήθηκαν Μπεϋζιανά μοντέλα παλινδρόμησης ώστε να ληφθεί υπόψη η χωρο-χρονική μετάδοση των μολυσματικών ζωνόσων αξιοποιώντας πραγματικά δεδομένα από επιδημίες. Αναπτύσσονται νέες μεθοδολογίες σε διαφορετικές κατευθύνσεις, συμπεριλαμβανομένης της ενσωμάτωσης της χωρικής πληροφορίας, την αντιμετώπιση της πολυσυγγραμμικότητας και τη χρήση κατάλληλων μικτών κατανομών. Επιπλέον, μία βασική συνιστώσα είναι η προσπάθεια τα μοντέλα να ενσωματωθούν στο πλαίσιο μιας κλαδωτής διαδικασίας από την οποία προκύπτει ο βέλτιστος τρόπος ελέγχου της νόσου.

Λέξεις-κλειδιά: Μπεϋζιανή μοντελοποίηση, επιδημικά μοντέλα, συνάρτηση πυρήνα.

Εισαγωγή

Ένας από τους σημαντικότερους στόχους στο πλαίσιο της έρευνας μιας επιδημικής έξαρσης είναι η αναπαράσταση του μηχανισμού μετάδοσής της, καθώς και των παραγόντων που επηρεάζουν τη δυναμική του μηχανισμού αυτού. Σε αυτό το πλαίσιο προτείνουμε μοντέλα παλινδρόμησης για τη περιγραφή της χωροχρονικής μετάδοσης μολυσματικών ζωνόσων, χρησιμοποιώντας δεδομένα από πραγματικές επιδημίες (ευλογιά του προβάτου) ώστε να επιλέξουμε προσεκτικά το καλύτερο - ανάμεσα σε υποψήφια - μοντέλο για την περιγραφή της δυναμικής της νόσου. Ως μελέτη περίπτωσης χρησιμοποιούμε τον Νομό Έβρου, κυρίως λόγω των ιδιαίτερων χαρακτηριστικών του (σύνορο της Ευρώπης με την Ασία) και της διαθεσιμότητας καλής ποιότητας τόσο ιστορικών όσο και τρεχόντων δεδομένων που συνδέονται με την εμφάνιση και εξάπλωση των ζωνόσων και τη δομή της κτηνοτροφίας στην περιοχή.

Μεθοδολογία

Το πλαίσιο της μοντελοποίησης που εφαρμόστηκε επεκτείνει παρόμοιες προσεγγίσεις (Choi et al., 2012; Branscum et al., 2008; Chhetri et al., 2010) με διάφορους τρόπους. Τα επιδημικά δεδομένα επιδεικνύουν συγκεκριμένα χαρακτηριστικά, όπως η χωρική φύση του μηχανισμού μετάδοσης της νόσου, η πολυσυγγραμμικότητα, η παρουσία υπερβολικών μηδενικών και ο περιβαλλοντικός θόρυβος. Προς αυτή την κατεύθυνση, ενσωματώ-

θηκε η διαθέσιμη χωρική πληροφορία (ακριβείς συντεταγμένες των μολυσμένων με ασθένεια κοπαδιών) με τη βοήθεια των συναρτήσεων πυρήνα (kernel functions) στα επιδημικά μοντέλα που προτείνονται. Προκύπτει ότι τα χωρικά μοντέλα περιγράφουν την εξάπλωση της νόσου πιο αποτελεσματικά από τα μη-χωρικά μοντέλα. Η επίδραση των διαφορετικών συναρτήσεων πυρήνα διερευνάται εκτενώς, πραγματοποιώντας μια συγκριτική ανάλυση μεταξύ των καταλληλότερων για αυτού του είδους τα δεδομένα (βλ. Πίνακα 1).

Επιπλέον, η υπάρχουσα μεθοδολογία βελτιώθηκε περαιτέρω αντιμετωπίζοντας το ζήτημα της πολυσυγγραμμικότητας. Στην ανάλυση κατά Bayes απαιτείται ο προσδιορισμός των εκ των προτέρων κατανομών. Σε αυτό το πλαίσιο χρησιμοποιήθηκε η κλίση των g-priors (Zellner, 1986) για την αντιμετώπιση των προβλημάτων πολυσυγγραμμικότητας μεταξύ των ανεξαρτήτων μεταβλητών που χρησιμοποιούνται, λαμβάνοντας υπόψη την απόκλιση των δεδομένων από την κανονική κατανομή (Bon and Held, 2011). Συγκεκριμένα διεξήχθη μια Μπεϋζιανή επιλογή μεταβλητών διερευνώντας την επίδραση των διαφόρων εκ των προτέρων κατανομών, όπως τεχνικές που βασίζονται σε παραλληλαγές από μίγματα όπως οι hyper-g priors (Liang et al., 2008). Η χρήση κατάλληλων μικτών κατανομών για τη μοντελοποίηση των επιπρόσθετων μηδενικών στα δεδομένα βελτιώνει σημαντικά τα μοντέλα σε σύγκριση με προηγούμενες μελέτες όπου γίνεται χρήση της κανονικής (Choi et al., 2012) ή της Poisson κατανομής (Branscum et al., 2008; Chhetri et al., 2010). Ελήφθη επίσης υπόψη η αυτοσυσχέτιση των δεδομένων ενσωματώνοντας στο μοντέλο κατάλληλες διαδικασίες (Ornstein-Uhlenbeck) για την μοντελοποίηση των συνεχούς χρόνου αυτοσυσχετίσεων (Struthers and McLeish, 2011). Τέλος, τα μοντέλα ενσωματώθηκαν στο πλαίσιο των κλαδωτών διαδικασιών (Andersson and Britton, 2000). Με τον τρόπο αυτό επιτυγχάνεται μία σύνδεση μεταξύ των παραδοσιακών χωρο-χρονικών μοντέλων με επιδημικές διαδικασίες, η οποία είναι ιδιαίτερα χρήσιμη στη λήψη αποφάσεων ελέγχου της νόσου.



Πίνακας 1. Σύνοψη των συναρτήσεων πυρήνα μετάδοσης που περιλαμβάνονται στα χωρο-χρονικά μοντέλα

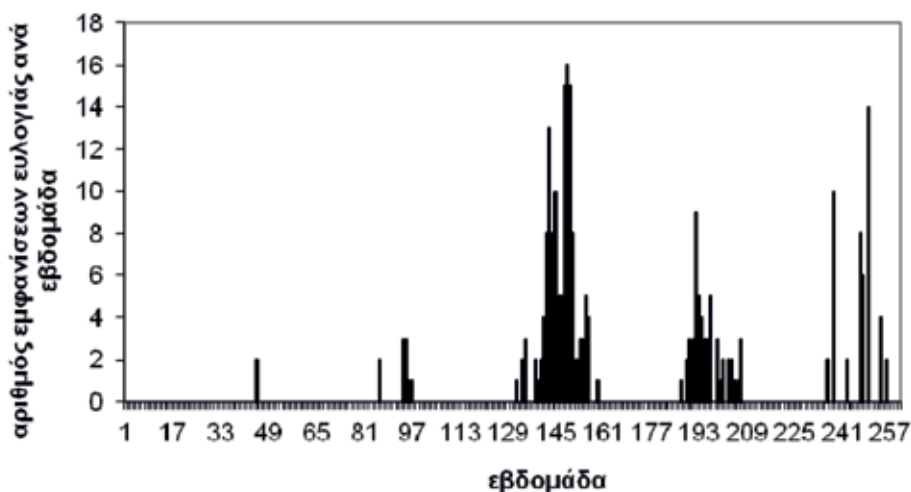
ΣΥΜΒΟΛΙΣΜΟΣ	$K(d_{kl}, \Theta_K)$	Θ_K	Αναφορά
A	$\left(1 + \frac{d_{kl}}{\alpha}\right)^{-c}$	(α, c)	Chis-Ster και Ferguson (2007)
B	$\exp\left\{-\left(\frac{d_{kl}}{\alpha}\right)^c\right\}$	(α, c)	Keeling και άλλοι (2001)
C	$\exp\left\{-\left(\frac{d_{kl}}{\alpha}\right)^c\right\} + r$	(α, c, r)	Diggle (2006)
D	$\alpha \exp\{-\alpha d_{kl}\}$	α	Szmaragd και άλλοι (2009)
E	$\frac{a}{\sqrt{\pi}} \exp\{-\alpha^2 d_{kl}^2\}$	α	Szmaragd και άλλοι (2009)
F	$\frac{a}{4} \exp\{-\alpha^{1/2} d_{kl}^{1/2}\}$	α	Szmaragd και άλλοι (2009)

Δεδομένα

Από το έτος 1995 μέχρι το έτος 2000 στην περιοχή της βορειο-ανατολικής Ελλάδας και συγκεκριμένα στην περιοχή του Έβρου υπήρξαν σημαντικές επιδημικές εξάρσεις ζωνόσων, κυρίως σχετικές με την ευλογιά των προβάτων. Η συγκεκριμένη νόσος έχει μεγάλη μεταδοτικότητα και προκαλεί σοβαρές διαταραχές στα ζώα. Η αντιμετώπιση είναι η θανάτωση των προσβεβλημένων και ύποπτων μόλυνσης κοπαδιών. Έχουμε συλλέξει όλα τα σχετικά ιστορικά στοιχεία για την ασθένεια, συγκεκριμένα εβδομαδιαίες μετρήσεις που σχετίζονται με τις φάρμες που προσβλήθηκαν και τις αντίστοιχες ημερομηνίες προσβολής και συντεταγμένες τους (Σχήμα 1). Επίσης χρησιμοποιούμε αντίστοιχα στοιχεία των μετεωρολογικών μεταβλητών τα οποία παραχωρήθηκαν από την Εθνική Μετεωρολογική Υπηρεσία.

Ανάλυση-Αποτελέσματα

Για την πραγματοποίηση της ανάλυσης χρησιμοποιήθηκε το στατιστικό πρόγραμμα WinBugs (Lunn et al., 2000) και τα μοντέλα μας προσαρμόστηκαν μέσω της προσομοίωσης Markov Chain Monte Carlo (MCMC) (Gelman et al., 2003). Η επιλογή ανάμεσα στα υποψήφια μοντέλα πραγματοποιήθηκε με τη χρήση της deviance, μιας γενίκευσης του R2. Η εμφάνιση της ασθένειας συν-



Σχήμα 1: Χρονική κατανομή των συμβάντων ευλογιάς του προβάτου στην περιοχή του Έβρου (1994-1998).

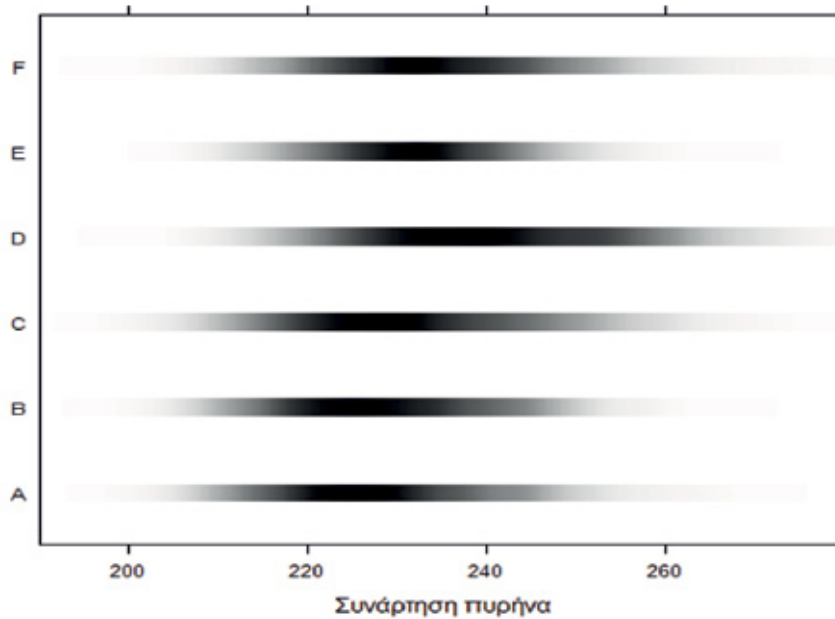
δέθηκε μέσα στο χώρο και το χρόνο με παράγοντες όπως η εποικότητα και περιβαλλοντικοί ή/και μετεωρολογικοί παράγοντες όπως τα επίπεδα θερμοκρασίας, υγρασίας και βροχής.

Τα αποτελέσματα της ανάλυσης έδειξαν ότι χρήσιμοι προγνωστικοί παράγοντες στην συχνότητα εμφάνισης της ευλογιάς των προβάτων βρέθηκαν να είναι συγκεκριμένες μετεωρολογικές μεταβλητές, με πιο σημαντικές τη μέση θερμοκρασία, τη μέγιστη θερμοκρασία και τα επίπεδα της υγρασίας (ενδημικός παράγων). Επίσης, η απόσταση ανάμεσα στις φάρμες βρέθηκε να είναι εξαιρετικά σημαντική (επιδημικός παράγοντας).

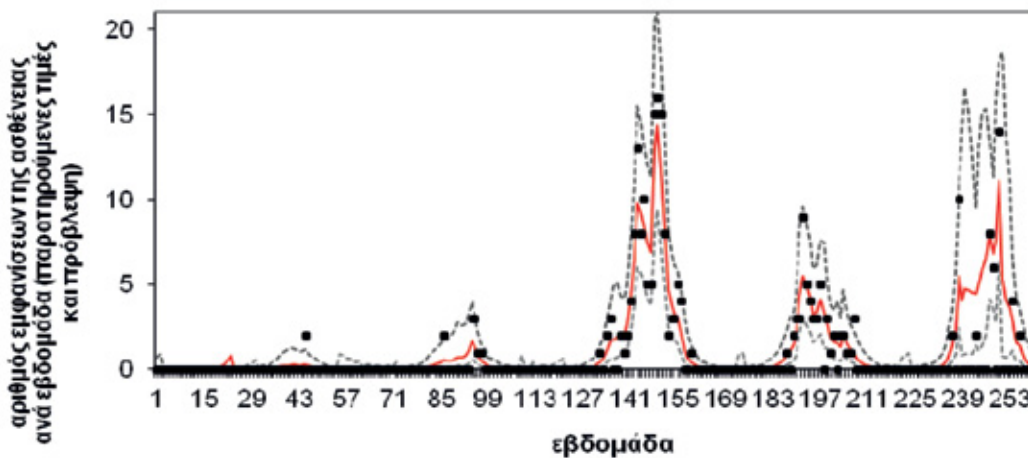
Η συγκριτική ανάλυση για την επιλογή της κατάλληλης συνάρτησης πυρήνα προτείνει τη χρήση πυρήνων

μακρίας ουράς (π.χ. πυρήνας A, σχήμα 2). Τέλος, το Σχήμα 3 απεικονίζει την προσαρμογή του καλύτερου

μοντέλου στα δεδομένα, υποδηλώνοντας ικανοποιητική προσαρμογή.



Σχήμα 2: Εκ των υστέρων κατανομή του κριτηρίου πληροφορίας (deviance) για τα χωρο-χρονικά μοντέλα βασισμένα στις 6 συναρτήσεις πυρήνα



Σχήμα 3: Σύγκριση εκτιμήσεων και πραγματικών τιμών του αριθμού των εμφανίσεων της ασθένειας για το επιλεγέν μοντέλο

Συμπεράσματα

Η μοντελοποίηση του μηχανισμού μετάδοσης μολυσματικών ζωνόσων αποτελεί ένα διεπιστημονικό πρόβλημα και στην παρούσα εργασία προτείνεται μία ενιαία μεθοδολογική προσέγγιση με στόχο την αντιμετώπιση ορισμένων από τα θεμελιώδη χαρακτηριστικά των επιδημικών δεδομένων. Οι συγγραφείς επικεντρώθηκαν στην εφαρμογή ενός ολιστικού πλαισίου για την μοντε-

λοποίηση των επιδημιολογικών δεδομένων προερχόμενων από ζωνόσους. Ειδικότερα η χρήση των μικτών κατανομών και η εισαγωγή κατάλληλων συναρτήσεων πυρήνα σε μοντέλα αυτού του είδους βελτίωσε σημαντικά την προσαρμογή τους, γεγονός που υποδηλώνει ότι οι μελλοντικές προσεγγίσεις μοντελοποίησης για την πρόβλεψη της διάδοσης των μολυσματικών ζωνόσων θα πρέπει να περιλαμβάνει παρόμοιες μεθοδολογίες.

- **Andersson, H. and Britton, T. (2000).** Stochastic epidemic models and their statistical analysis. Springer Lecture Notes in Statistics, 151. Springer-Verlag, New York.
- **Bovi, D.S. and Held, L. (2011).** Hyper-g priors for generalized linear models. *Bayesian Analysis*, 6(3), 387–410.
- **Branscum, A.J., Perez, A.M., Johnson, W.O. and Thurmond, M.C. (2008).** Bayesian spatiotemporal analysis of foot-and-mouth disease data from the Republic of Turkey. *Epidemiology and Infection*, 136, 833-842.
- **Chhetri, B.K., Perez, A.M. and Thurmond, M.C. (2010).** Factors associated with spatial clustering of foot-and-mouth disease in Nepal. *Tropical Animal Health and Production*, 42, 1441-1449.
- **Chis-Ster, I.C. and Ferguson, N.M. (2007).** Transmission parameters of the 2001 Foot and Mouth epidemic in Great Britain. *PLoS ONE*, 6, e502.
- **Choi, Y.K., Johnson, W.O., Jones, G., Perez, A. and Thurmond, M.C. (2012).** Modelling and predicting temporal frequency of foot-and-mouth disease cases in countries with endemic foot-and-mouth disease. *Journal of the Royal Statistical Society A*, 175(2), 619-636.
- **Diggle, P.J. (2006).** Spatio-temporal point processes, partial likelihood, foot and mouth disease. *Statistical Methods in Medical Research*, 15, 325-336.
- **Gelman A, Meng XL, Stern HS, Rubin DB. (2003).** *Bayesian Data Analysis*, (2nd ed), Chapman and Hall, London.
- **Keeling, MJ, Woolhouse, MEJ, Shaw, DJ, Matthews, L., Chase-Topping, M., Haydon, DT., Cornell, SJ., Kappey, J., Wilesmith, J., Grenfell, BT. (2001).** Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. *Science*, 294, 813-817.
- **Lambert, D. (1992).** Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34, 1–14.
- **Liang, F., Paulo, R., Molina, G., Clyde, M.A. and Berger, J.O. (2008).** Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association*, 103(481), 410-423.
- **Lunn, D.J., Thomas, A., Best, N. and Spiegelhalter, D. (2000).** WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*, 10, 325–337.
- **Paul, M. and Held, L. (2011).** Predictive assessment of a non-linear random effects model for multivariate time series of infectious disease counts. *Statistics in Medicine*, 30, 1118-1136.
- **Struthers, C.A. and McLeish, D.L. (2011).** A particular diffusion model for incomplete longitudinal data: application to the multicenter AIDS cohort study. *BMC Biostatistics*, 12(3), 493-505.
- **Szmaragd, C., Wilson, A.J, Carpenter, S., Wood, J.L.N, Mellor, P.S. and Gubbins, S. (2009).** A modeling framework to describe the transmission of bluetongue virus within and between farms in Great Britain. *PLoS ONE*, 4(11), e7741.
- **Zellner, A. (1986).** On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In *Bayesian inference and decision techniques: Essays in honor of Bruno de Finetti*, eds. P.K. Goel and A. Zellner, Amsterdam: North-Holland/Elsevier, 233-243.