

ΕΚΤΙΜΗΣΗ ΤΗΣ ΑΒΕΒΑΙΟΤΗΤΑΣ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ ΚΑΥΣΙΜΩΝ

Δημήτριος Θεοδώρου, Φανούριος Ζαννίκος

Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ), Σχολή Χημικών Μηχανικών, Εργαστήριο
Τεχνολογίας Καυσίμων και Λιπαντικών
Πολυτεχνειούπολη, Ηρώων Πολυτεχνείου 9, 15773 Ζωγράφου, Αθήνα
e-mail: dtheodoro@teemail.gr

Περίληψη

Η δειγματοληψία αποτελεί ένα βασικό στάδιο των μετρητικών διαδικασιών και συνεισφέρει σημαντικά στην αβεβαιότητα των εργαστηριακών μετρήσεων. Μια αξιόπιστη εκτίμηση της αβεβαιότητας λόγω δειγματοληψίας μπορεί να οδηγήσει σε καλύτερο έλεγχο των κινδύνων που συνδέονται με αποφάσεις περί συμμόρφωσης ή μη, ενός καυσίμου με προδιαγραφές που επιβάλλει η νομοθεσία. Στην παρούσα εργασία περιγράφονται και συγκρίνονται ως προς τα αποτελέσματά τους, τρεις εμπειρικές στατιστικές μεθοδολογίες («κλασική» ANOVA, ανθεκτική ANOVA και στατιστική εύρους τιμών) χρησιμοποιώντας δεδομένα ενός ισορροπημένου πειραματικού σχεδίου (balanced experimental design) το οποίο περιλαμβάνει διπλές δειγματοληψίες από 104 πρατήρια καυσίμων και στη συνέχεια εις διπλούν αναλύσεις του κάθε δείγματος.

Οι τρεις μεθοδολογίες χρησιμοποιούνται για την εκτίμηση της αβεβαιότητας λόγω δειγματοληψίας καυσίμου (ντίζελ κίνησης) και λόγω της αναλυτικής διαδικασίας προσδιορισμού περιεκτικότητας σε θείο. Τα αποτελέσματα των τριών μεθοδολογιών διαφέρουν στατιστικά με την διευρυμένη αβεβαιότητα (επίπεδο εμπιστοσύνης 95%) λόγω δειγματοληψίας να κυμαίνεται από 0,34 έως 0,40 mg kg⁻¹ (ή από 4,4 έως 5,1 % εκφρασμένη ως σχετική διευρυμένη αβεβαιότητα) ανάλογα με την μεθοδολογία που ακολουθήθηκε. Η εκτίμηση της ανθεκτικής ANOVA για την αβεβαιότητα δειγματοληψίας (0,34 mg kg⁻¹ ή 4,8%) κρίνεται ως περισσότερο αξιόπιστη λόγω της παρουσίας ακραίων (έκτοπων) τιμών στις 104 ομάδες δεδομένων που χρησιμοποιήθηκαν στους υπολογισμούς. Η ανθεκτική ANOVA σε αντίθεση με την «κλασική» ANOVA και τη στατιστική εύρους τιμών είναι σε θέση να ανιχνεύσει και να αντιμετωπίσει τις ακραίες τιμές, περιορίζοντας την επίδρασή τους στα εκτιμώμενα μεγέθη.

Επιπλέον, τα αποτελέσματα της παρούσας εργασίας καταδεικνύουν ότι στην περίπτωση δειγματοληψίας καυσίμων από πρατήρια καυσίμων η κύρια συνεισφορά στη αβεβαιότητα της μέτρησης προέρχεται από τη διαδικασία της εργαστηριακής ανάλυσης με την αβεβαιότητα της δειγματοληψίας να περιορίζεται στο 29% της συνολικής αβεβαιότητας.

Λέξεις-Κλειδιά: Αβεβαιότητα, Δειγματοληψία καυσίμων, ANOVA, Στατιστική εύρους τιμών

Abstract

Sampling is an important part of any measurement process and is therefore recognized as an important contributor to the measurement uncertainty. A reliable estimation of the uncertainty arising from sampling of fuels leads to a better control of risks associated with decisions concerning whether product specifications are met or not. The present work describes and compares the results of three empirical statistical methodologies (classical ANOVA, robust ANOVA and range statistics) using data from a balanced experimental design, which includes duplicate samples analyzed in duplicate from 104 sampling targets (gas stations). These methodologies are used for the estimation of uncertainty arising from the manual sampling of fuel (automotive diesel) and the subsequent sulfur mass content determination. The results of the three methodologies statistically differ, with the expanded uncertainty of sampling being in the range of 0.34 – 0.40 mg kg⁻¹, while the relative expanded uncertainty lying in the range of 4.8 - 5.1%, depending on the methodology used. The estimation of robust ANOVA (sampling expanded uncertainty of 0.34 mg kg⁻¹ or 4.8% in relative terms) is considered more reliable, because of the presence of outliers within the 104 datasets used for the calculations. Robust ANOVA, in contrast to classical ANOVA and range statistics, accommodates outlying values, lessening their effects on the produced estimates. The results of this work also show that, in the case of manual sampling of fuels, the main contributor to the whole measurement uncertainty is the analytical measurement uncertainty, with the sampling uncertainty accounting only for the 29% of the total uncertainty.

Keywords: Uncertainty, Fuel sampling, ANOVA, Range statistics

1. Εισαγωγή

Ο σκοπός μιας δειγματοληψίας είναι να ληφθεί ένα μικρό τμήμα ενός υλικού (δείγμα) από ένα επιλεγμένο σύστημα (στόχος δειγματοληψίας) με τέτοιο τρόπο ώστε να εξασφαλίζεται ότι το δείγμα είναι αντιπροσωπευτικό της σύνθεσης του στόχου δειγματοληψίας (Heydorn et al. 2004). Αντιπροσωπευτικά δείγματα πετρελαίου και προϊόντων πετρελαίου απαιτούνται για τον προσδιορισμό συγκεκριμένων χημικών και φυσικών ιδιοτήτων, οι οποίες συχνά χρησιμοποιούνται για τη αξιολόγηση της συμμόρφωσης των προϊόντων αυτών με εμπορικές και νομοθετικές προδιαγραφές. Η δειγματοληψία έχει ιδιαίτερη σημασία όταν απαιτείται να εκτιμήσει κανείς την αβεβαιότητα της μέτρησης, όπου θα πρέπει να συνυπολογίζει και τυχόν αβεβαιότητα λόγω δειγματοληψίας (ASTM 2012a). Η αβεβαιότητα δειγματοληψίας ορίζεται ως το μέρος εκείνο της συνολικής αβεβαιότητας της μέτρησης που οφείλεται στη δειγματοληψία. Οι βασικές αρχές και οι διαδικασίες για την εκτίμηση της αβεβαιότητας των μετρήσεων που προκύπτουν από τη δειγματοληψία περιγράφονται στην Οδηγία που δημοσιεύθηκε από τη Eurachem (2007), καθώς και στο Εγχειρίδιο της Nordtest (2007).

Υπάρχουν δύο γενικές προσεγγίσεις-μέθοδοι για την εκτίμηση της αβεβαιότητας, η μέθοδος της μοντελοποίησης και η εμπειρική μέθοδος (Ramsey 1998, Ramsey 2004, Lyn et al. 2007). Μία από τις πιο συχνά χρησιμοποιούμενες εμπειρικές μεθόδους είναι η μέθοδος των εις διπλούν μετρήσεων βάσει ενός ισορροπημένου πειραματικού σχεδίου. Αυτή η μέθοδος περιλαμβάνει τη λήψη διπλών δειγμάτων από τους στόχους της

Δημήτριος Θεοδώρου, Φανούριος Ζαννίκος, Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ), Σχολή Χημικών Μηχανικών, Εργαστήριο Τεχνολογίας Καυσίμων και Λιπαντικών.

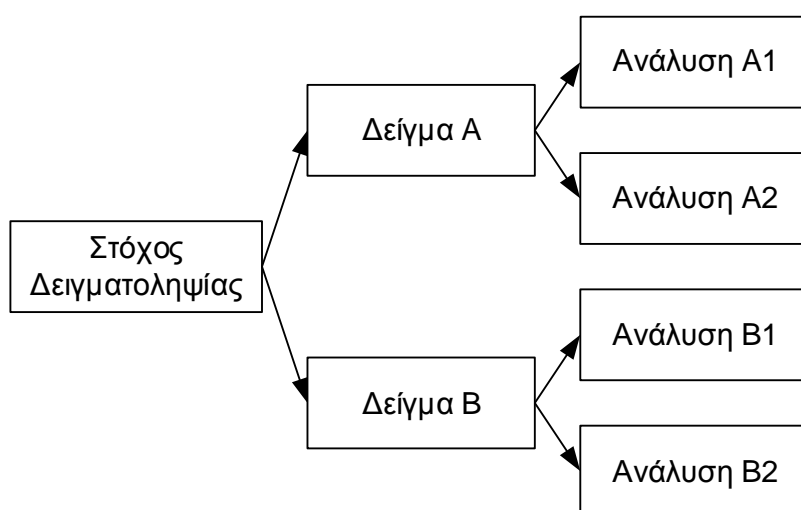
Εκτίμηση της αβεβαιότητας της δειγματοληψίας καυσίμων

δειγματοληψίας εφαρμόζοντας το ίδιο πρωτόκολλο δειγματοληψίας και την ανάλυση των δειγμάτων εις διπλούν υπό συνθήκες επαναληψιμότητας. Μετά από κατάλληλη στατιστική επεξεργασία των παραγόμενων δεδομένων είναι δυνατό να εκτιμηθεί η αβεβαιότητα της δειγματοληψίας.

Ο σκοπός της παρούσας εργασίας είναι να παρουσιάσει και να συγκρίνει τρεις στατιστικές μεθοδολογίες που χρησιμοποιούνται για την εκτίμηση της αβεβαιότητας που οφείλεται στη δειγματοληψία των καυσίμων από πρατήρια, χρησιμοποιώντας τα αποτελέσματα αναλύσεων από εις διπλούν δειγματοληψίες.

2. Πρωτόκολλο δειγματοληψίας και σχεδιασμός μετρήσεων

Υλοποιήθηκε ένα ισορροπημένο πειραματικό σχέδιο. Διπλά δείγματα ελήφθησαν από 104 πρατήρια υγρών καυσίμων, τα οποία επιλέχθηκαν τυχαία και αποτελούσαν το 10,9% των 950 πρατηρίων τα οποία παρακολουθούνται από το εργαστήριο. Το σύστημα δειγματοληψίας – ανάλυσης που χρησιμοποιήθηκε παρουσιάζεται στο Σχήμα 1. Τα διπλά δείγματα ελήφθησαν επαναλαμβάνοντας το ίδιο πρωτόκολλο δειγματοληψίας. Το πρωτόκολλο δειγματοληψίας που χρησιμοποιήθηκε ήταν βασισμένο στην πρότυπη μέθοδο ASTM D 4057 (2012a) που περιγράφει μεθοδολογίες μη αυτόματης δειγματοληψίας πετρελαίου και προϊόντων πετρελαίου.



Σχήμα 1: Ισορροπημένο πειραματικό σχέδιο για την εκτίμηση της αβεβαιότητας της δειγματοληψίας

3. Αναλυτική Μεθοδολογία

Ο προσδιορισμός της περιεκτικότητας των δειγμάτων νίζελ κίνησης σε θείο διεξήχθη στο Εργαστήριο Τεχνολογίας Καυσίμων και Λιπαντικών του ΕΜΠ, το οποίο εφαρμόζει σύστημα διαχείρισης ποιότητας κατά ISO 9001 και ISO / IEC 17025 και συμμετέχει περιοδικά σε διεργαστηριακά προγράμματα δοκιμών ικανότητας για μεγάλο εύρος αναλύσεων καυσίμων. Τα διπλά δείγματα αναλύθηκαν εις διπλούν κάτω από συνθήκες επαναληψιμότητας. Για τον προσδιορισμό της περιεκτικότητας σε θείο χρησιμοποιήθηκε ο αναλυτής θείου ANTEK 9000S, εξοπλισμένος με αυτόματο δειγματολήπτη. Ο

Δημήτριος Θεοδώρου, Φανούριος Ζαννίκος, Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ), Σχολή Χημικών Μηχανικών, Εργαστήριο Τεχνολογίας Καυσίμων και Λιπαντικών.

Εκτίμηση της αβεβαιότητας της δειγματοληψίας καυσίμων

εξοπλισμός αυτός συμμορφώνεται πλήρως με τα πρότυπα ASTM D 5453 (2012b) και ISO 20846 (2011) για τον προσδιορισμό θείου σε καύσιμα αυτοκινήτων. Η πλειοψηφία των αποτελεσμάτων των μετρήσεων των δειγμάτων βρέθηκε να περιέχει θείο λιγότερο από ή πολύ κοντά στα 10 mg kg^{-1} , το οποίο είναι το νομοθετικό (άνω) όριο της ΕΕ για την περιεκτικότητα του ντίζελ κίνησης σε θείο.

4. Μεθοδολογίες ανάλυσης δεδομένων

4.1 Στατιστικό μοντέλο και εκτίμηση αβεβαιοτήτων

Το στατιστικό μοντέλο που περιγράφει τη σχέση μεταξύ μετρούμενης και πραγματικής τιμής της συγκέντρωσης μιας ουσίας, βάσει μια μέτρησης, x , έχει την ακόλουθη μορφή:

$$x = X_{\text{true}} + \varepsilon_{\text{sampling}} + \varepsilon_{\text{analysis}} \quad (1)$$

όπου X_{true} , είναι η πραγματική μέση συγκέντρωση της ουσίας στο σημείο δειγματοληψίας, $\varepsilon_{\text{sampling}}$, είναι το συνολικό σφάλμα λόγω δειγματοληψίας, που εκφράζεται από μια διακύμανση $\sigma_{\text{sampling}}^2$ και $\varepsilon_{\text{analysis}}$, είναι το συνολικό αναλυτικό σφάλμα, που εκφράζεται από μια διακύμανση, $\sigma_{\text{analysis}}^2$ (Eurachem 2007). Αν οι πηγές της διακύμανσης είναι ανεξάρτητες, τότε η συνολική διακύμανση της μέτρησης, $\sigma_{\text{measurement}}^2$, για ένα σημείο δειγματοληψίας δίνεται από τη σχέση:

$$\sigma_{\text{measurement}}^2 = \sigma_{\text{sampling}}^2 + \sigma_{\text{analysis}}^2 \quad (2)$$

Αν χρησιμοποιηθούν οι στατιστικές εκτιμήτριες, s^2 , για την προσέγγιση των διακυμάνσεων, σ^2 , τότε η εξίσωση (2) γίνεται:

$$s_{\text{measurement}}^2 = s_{\text{sampling}}^2 + s_{\text{analysis}}^2 \quad (3)$$

Τα συστατικά της διακύμανσης της μέτρησης, οι διακυμάνσεις της δειγματοληψίας και της ανάλυσης, που αντιπροσωπεύουν τις αντίστοιχες αβεβαιότητες, μπορούν να διαχωριστούν και να εκτιμηθούν χρησιμοποιώντας κατάλληλες στατιστικές μεθόδους, όπως αυτές που παρουσιάζονται στην παρούσα εργασία.

Για να υπολογιστεί ένα διάστημα κάλυψης (διευρυμένη αβεβαιότητα) που αντιστοιχεί σε πιθανότητα κάλυψης περίπου 95%, οι τυπικές αποκλίσεις, s , πολλαπλασιάζονται με ένα συντελεστή κάλυψης που ισούται με δύο (2). Η διευρυμένη αβεβαιότητα μέτρησης, δειγματοληψίας και ανάλυσης, $U_{\text{measurement}}$, U_{sampling} και U_{analysis} , αντίστοιχα, υπολογίζονται σύμφωνα με τις ακόλουθες εξισώσεις:

$$U_{\text{measurement}} = 2s_{\text{measurement}} \quad (4)$$

$$U_{\text{sampling}} = 2s_{\text{sampling}} \quad (5)$$

$$U_{\text{analysis}} = 2s_{\text{analysis}} \quad (6)$$

Η διευρυμένη αβεβαιότητα μπορεί επίσης να εκφραστεί σε σχέση με την αναφερόμενη τιμή x (ως ποσοστό), ως σχετική διευρυμένη αβεβαιότητα U (%):

$$U_{\text{measurement}} (\%) = 100 \frac{2s_{\text{measurement}}}{x} \% \quad (7)$$

$$U_{\text{sampling}} (\%) = 100 \frac{2s_{\text{sampling}}}{x} \% \quad (8)$$

$$U_{\text{analysis}} (\%) = 100 \frac{2s_{\text{analysis}}}{x} \% \quad (9)$$

4.2 Εκτίμηση αβεβαιοτήτων με «κλασική» ANOVA

Η ANOVA (ανάλυση διακύμανσης) είναι μία στατιστική τεχνική με την οποία οι διακυμάνσεις που προέρχονται από διαφορετικές πηγές μπορούν να απομονωθούν και να εκτιμηθούν (Ellison et al. 2010). Ο πιο απλός τύπος ANOVA είναι η ANOVA κατά ένα παράγοντα (one way), η οποία εξετάζει μία ανεξάρτητη μεταβλητή και μία εξαρτημένη μεταβλητή. Η ANOVA μπορεί να εφαρμοστεί στα δεδομένα που παράγονται από την εφαρμογή ενός ισορροπημένου πειραματικού σχεδίου, προκειμένου να εκτιμηθεί η αβεβαιότητα δειγματοληψίας. Οι εκτιμήσεις της ANOVA βασίζονται στις διαφορές από τις μέσες τιμές, και όχι σε εύρη όπως ισχύει στην στατιστική εύρους τιμών. Ο Πίνακας 1 παρουσιάζει τις μαθηματικές σχέσεις της ANOVA κατά ένα παράγοντα που χρησιμοποιούνται για την εκτίμηση της αβεβαιότητας της δειγματοληψίας και της ανάλυσης.

Πίνακας 1: Εκτίμηση των αβεβαιοτήτων δειγματοληψίας και ανάλυσης με χρήση ANOVA κατά ένα παράγοντα, χρησιμοποιώντας δεδομένα μετρήσεων ενός ισορροπημένου πειραματικού σχεδίου με n σημεία δειγματοληψίας ($i=1,2,\dots,n$), 2 δείγματα ($j=A, B$) από κάθε σημείο και 2 αναλύσεις ($k=1, 2$) για κάθε δείγμα.

	Πηγή της διακύμανσης	
	Δειγματοληψία	Ανάλυση
Απόσταση από τη μέση τιμή (D) ^{a,b,c}	$D_{i(x)} = \bar{X}_i - \bar{x}_{iA} = \bar{X}_i - \bar{x}_{iB} $	$D_{iA(x)} = x_{iA1} - \bar{x}_{iA} = x_{iA2} - \bar{x}_{iA} $ $D_{iB(x)} = x_{iB1} - \bar{x}_{iB} = x_{iB2} - \bar{x}_{iB} $
Άθροισμα τετραγώνων (SS)	$SS_{\text{sampling}} = 4 \sum_i D_{i(x)}^2$	$SS_{\text{analysis}} = 2 \sum_i (D_{iA(x)}^2 + D_{iB(x)}^2)$
Βαθμοί ελευθερίας (df)	$2n - n = n$	$2 \cdot 2n - 2n = 2n$
Μέσα τετράγωνα (MS)	$MS_{\text{sampling}} = \frac{SS_{\text{sampling}}}{df_{\text{sampling}}}$	$MS_{\text{analysis}} = \frac{SS_{\text{analysis}}}{df_{\text{analysis}}}$
Διακύμανση (V)	$V_{\text{sampling}} = \frac{MS_{\text{sampling}} - MS_{\text{analysis}}}{2}$	$V_{\text{analysis}} = MS_{\text{analysis}}$
Τυπική αβεβαιότητα (s)	$s_{\text{sampling}} = \sqrt{V_{\text{sampling}}}$	$s_{\text{analysis}} = \sqrt{V_{\text{analysis}}}$

^a \bar{X}_i : μέση τιμή του σημείου δειγματοληψίας i (2 δείγματα – 4 αναλύσεις)

^b \bar{x}_{ij} : μέση τιμή των 2 αναλύσεων του δείγματος j (A ή B) του σημείου δειγματοληψίας i

^c x_{ijk} : μετρηθείσα τιμή από το σημείο δειγματοληψίας i , για το δείγμα j και για την ανάλυση k (1 ή 2)

4.3 Εκτίμηση αβεβαιότητας με ανθεκτική ANOVA

Η εγκυρότητα των αποτελεσμάτων της «κλασικής» ANOVA βασίζεται σε τρεις παραδοχές: (α) οι διακυμάνσεις θα πρέπει να είναι ανεξάρτητες, (β) κάθε επίπεδο της διακύμανσης θα πρέπει να είναι ομοιογενές, και να μη μεταβάλλεται συστηματικά εντός του επιπέδου και (γ) η κατανομή των σφαλμάτων σε κάθε επίπεδο διακύμανσης πρέπει να προσεγγίζει την κανονική κατανομή (Ramsey 1992). Η εφαρμογή της ανθεκτικής ANOVA που χρησιμοποιεί ανθεκτικά στατιστικά μεγέθη έχει αποδειχθεί ότι είναι ιδιαίτερα κατάλληλη για την εκτίμηση διακυμάνσεων, σε περιπτώσεις όπου η εγκυρότητα της «κλασικής» ANOVA είναι αμφισβητήσιμη (Eurachem 2007). Η ανθεκτική ANOVA χρησιμοποιεί ανθεκτικές εκτιμήσεις του μέσου όρου και της τυπικής απόκλισης που υπολογίζονται από μια επαναληπτική διαδικασία (Royal Society of Chemistry 1989, Ramsey 1992). Στην παρούσα εργασία, η ανθεκτική ANOVA υλοποιήθηκε χρησιμοποιώντας το εξειδικευμένο λογισμικό, Roban.exe, που αναπτύχθηκε από την Analytical Methods Committee (AMC), στη Μεγάλη Βρετανία (Royal Society of Chemistry 2001).

4.4 Εκτίμηση αβεβαιότητας με χρήση στατιστικής εύρους τιμών

Η στατιστική εύρους τιμών μπορεί επίσης να χρησιμοποιηθεί για τον υπολογισμό των τυπικών αποκλίσεων μέσω της επεξεργασίας των δεδομένων ενός ισορροπημένου πειραματικού σχεδίου (Nordtest 2007, ISO 2002). Η στατιστική εύρους τιμών, όπως και η «κλασική» ANOVA κάνει τη παραδοχή ύπαρξης κανονικής κατανομής ενώ οι υπολογισμοί γίνονται χρησιμοποιώντας διαφορές μεταξύ διπλών μετρήσεων. Στην πραγματικότητα η διακύμανση της δειγματοληψίας υπολογίζεται έμμεσα ως η διαφορά των διακυμάνσεων της μέτρησης και της ανάλυσης. Ο Πίνακας 2 παρουσιάζει τις μαθηματικές σχέσεις της στατιστικής εύρους τιμών που χρησιμοποιούνται για την εκτίμηση της αβεβαιότητας δειγματοληψίας και ανάλυσης.

Πίνακας 2: Εκτίμηση των αβεβαιότητων δειγματοληψίας και ανάλυσης με χρήση στατιστικής εύρους τιμών χρησιμοποιώντας δεδομένα μετρήσεων ενός ισορροπημένου πειραματικού σχεδίου με n σημεία δειγματοληψίας ($i=1,2,\dots,n$), 2 δείγματα ($j=A, B$) από κάθε σημείο και 2 αναλύσεις ($k=1, 2$) για κάθε δείγμα.

Παράμετρος	Εξίσωση υπολογισμού
Διαφορές εις διπλούν αναλύσεων δείγματος A (D_{iA}) ^a	$D_{iA} = x_{iA1} - x_{iA2} $
Διαφορές εις διπλούν αναλύσεων δείγματος B (D_{iB}) ^a	$D_{iB} = x_{iB1} - x_{iB2} $
Διαφορές των μέσων τιμών των δύο μετρήσεων (D_i) ^b	$D_i = \bar{x}_{iA} - \bar{x}_{iB} $
Μέσο εύρος της μέτρησης ($D_{\text{measurement}}$)	$D_{\text{measurement}} = \frac{\sum_i D_i}{n}$

Δημήτριος Θεοδώρου, Φανούριος Ζαννίκος, Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ), Σχολή Χημικών Μηχανικών, Εργαστήριο Τεχνολογίας Καυσίμων και Λιπαντικών.

Εκτίμηση της αβεβαιότητας της δειγματοληψίας καυσίμων

Παράμετρος	Εξίσωση υπολογισμού
Μέσο εύρος των αναλύσεων (D_{analysis})	$D_{\text{analysis}} = \frac{1}{2} \left(\frac{\sum_i D_{iA}}{n} + \frac{\sum_i D_{iB}}{n} \right)$
Τυπική απόκλιση της ανάλυσης (s_{analysis})	$s_{\text{analysis}} = \frac{D_{\text{analysis}}}{1.128}$
Τυπική απόκλιση της μέτρησης ($s_{\text{measurement}}$)	$s_{\text{measurement}} = \frac{D_{\text{measurement}}}{1.128}$
Τυπική απόκλιση της δειγματοληψίας (s_{sampling})	$s_{\text{sampling}} = \sqrt{s_{\text{measurement}}^2 - \left(\frac{s_{\text{analysis}}}{\sqrt{2}} \right)^2}$ <p>Η s_{analysis} διαιρείται με τετρ. ρίζα του 2 διότι το αποτέλεσμα της ανάλυσης του κάθε δείγματος προκύπτει ως μέσος όρος 2 μετρήσεων – τυπική απόκλιση μέσης τιμής</p>

a x_{ijk} : μετρηθείσα τιμή από το σημείο δειγματοληψίας i , για το δείγμα j και για την ανάλυση k (1 ή 2)

b \bar{x}_{ij} : μέση τιμή των 2 αναλύσεων του δείγματος j (A ή B) του σημείου δειγματοληψίας i

5. Αποτελέσματα

Τα αποτελέσματα της επεξεργασίας των μετρήσεων (εις διπλούν αναλύσεις 104 διπλών δειγμάτων) με τη χρήση των τριών στατιστικών μεθοδολογιών παρουσιάζονται στον Πίνακα 3 και στο Σχήμα 3. Η διευρυμένη αβεβαιότητα της δειγματοληψίας κυμαίνεται από 0,34 έως 0,40 mg kg⁻¹, ενώ η σχετική διευρυμένη αβεβαιότητα από 4,8 έως 5,1%, ανάλογα με τη στατιστική μεθοδολογία που χρησιμοποιήθηκε.

Μεταξύ των δεδομένων υπάρχουν επτά ακραίες (έκτοπες) τιμές μεταξύ σημείων δειγματοληψίας, οι οποίες επιβεβαιώθηκαν ως τέτοιες εφαρμόζοντας το τεστ Grubbs. Επιπλέον, τρεις ομάδες δεδομένων αναγνωρίστηκαν ως ακραίες τιμές λόγω ανάλυσης ή δειγματοληψίας εφαρμόζοντας το τεστ Cochran (Σχήμα 2). Συνολικά, 9 από τις 104 ομάδες δεδομένων (8,7%) αναγνωρίστηκαν στατιστικά ως ακραίες τιμές. Ως εκ τούτου, τα αποτελέσματα της ανθεκτικής ANOVA, η οποία δεν επηρεάζεται από μικρό αριθμό ακραίων τιμών (λιγότερο από 10%), μπορεί να θεωρηθούν ως περισσότερο αξιόπιστα σε σύγκριση με τα αποτελέσματα της Στατιστικής Εύρους και της «κλασικής» ANOVA.

Sampling Target	Analysis A1	Analysis A2	Analysis B1	Analysis B2	
.....	
84	7,61	7,88	7,95	7,74	
85	8,71	8,22	8,19	8,39	
86	32,99	31,04	33,93	28,89	between target outlier (Grubbs test)
87	30,96	31,06	30,79	30,43	both between target outlier (Grubbs test) and sampling/ analytical outlier (Cochran test)
88	7,49	7,55	7,49	7,58	
89	7,35	6,65	8,01	7,29	
90	8,17	8,09	8,99	8,95	
91	9,44	8,64	10,44	9,42	
92	5,94	5,87	7,13	6,78	
93	7,72	7,38	7,57	7,11	
94	7,39	6,87	7,18	7,29	
95	7,85	8,40	7,61	7,78	
96	6,86	6,65	6,92	7,27	
97	8,13	7,97	8,73	8,03	
98	8,18	8,20	8,40	7,91	
99	7,29	8,01	7,87	7,91	
100	11,97	11,16	11,69	11,86	
101	10,73	9,90	8,06	8,68	sampling/ analytical outlier (Cochran test)
102	7,55	7,22	7,09	7,17	
103	9,67	10,41	8,57	7,49	sampling/ analytical outlier (Cochran test)
104	7,57	7,60	8,08	7,80	

Σχήμα 2. Μέρος των πειραματικών δεδομένων όπου απεικονίζονται ενδεικτικές ομάδες δεδομένων που αναγνωρίστηκαν ως ακραίες τιμές μέσω των τεστ Cochran και Grubbs.

Είναι φανερό από τα αποτελέσματα της ανθεκτικής ANOVA ότι η αβεβαιότητα της μέτρησης προέρχεται κυρίως από την διακύμανση της ανάλυσης. Συγκεκριμένα, η αβεβαιότητα της ανάλυσης αντιπροσωπεύει το 71% της συνολικής αβεβαιότητας της μέτρησης. Αυτό αφήνει «χώρο» για την αποτελεσματική μείωση της συνολικής αβεβαιότητας, η οποία μπορεί να επιτευχθεί κάνοντας περισσότερες μετρήσεις και υπολογίζοντας τη μέση τους, αντί για τη διενέργεια μιας μονής μέτρησης. Σε αυτή την περίπτωση, η τυπική απόκλιση της μέσης τιμής ελαττώνεται οδηγώντας σε μικρότερες συνεισφορές στην αβεβαιότητα λόγω τυχαίου σφάλματος.

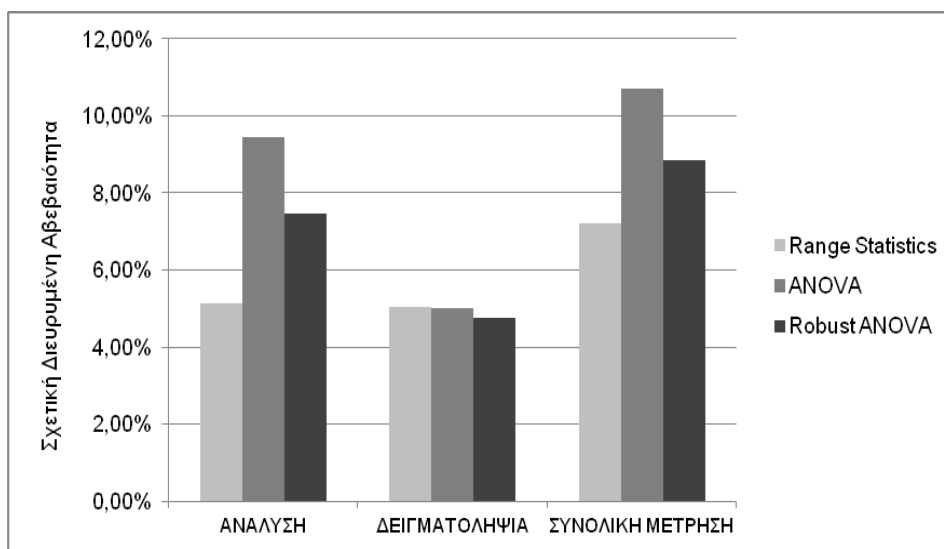
Πίνακας 3: Αποτελέσματα υπολογισμών με χρήση στατιστικής εύρους τιμών, «κλασικής» ANOVA και ανθεκτικής ANOVA. Όλες οι αβεβαιότητες U έχουν εκτιμηθεί με χρήση συντελεστή κάλυψης ίσο με 2 που αντιστοιχεί σε επίπεδο εμπιστοσύνης 95% περίπου.

	Στατιστική Εύρους Τιμών	«Κλασική» ANOVA	Ανθεκτική ANOVA
Μέση τιμή (mg kg ⁻¹)	7,988	7,988	7,079
$s_{analysis}$ (mg kg ⁻¹)	0,205	0,378	0,265
$s_{sampling}$ (mg kg ⁻¹)	0,202	0,200	0,169
$s_{measurement}$ (mg kg ⁻¹)	0,288	0,427	0,314
$U_{analysis}$ (mg kg ⁻¹)	0,411	0,755	0,529
$U_{analysis}$ (%)	5,1	9,5	7,5
$U_{sampling}$ (mg kg ⁻¹)	0,404	0,401	0,337
$U_{sampling}$ (%)	5,1	5,0	4,8
$U_{measurement}$ (mg kg ⁻¹)	0,576	0,855	0,628
$U_{measurement}$ (%)	7,2	10,7	8,9
Συνεισφορά στην αβεβαιότητα λόγω ανάλυσης (%)	51	78	71
Συνεισφορά στην αβεβαιότητα λόγω δειγματοληψίας (%)	49	22	29

Δημήτριος Θεοδώρου, Φανούριος Ζαννίκος, Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ), Σχολή Χημικών Μηχανικών, Εργαστήριο Τεχνολογίας Καυσίμων και Λιπαντικών.

Εκτίμηση της αβεβαιότητας της δειγματοληψίας καυσίμων

5ο Τακτικό Εθνικό Συνέδριο Μετρολογίας, Εθνικό Ίδρυμα Ερευνών, Αθήνα, 9-10 Μαΐου 2014



Σχήμα 3: Σύγκριση των σχετικών διευρυμένων αβεβαιοτήτων όπως αυτές εκτιμήθηκαν με χρήση στατιστικής εύρους τιμών, «κλασικής» ANOVA και ανθεκτικής ANOVA. Όλες οι αβεβαιότητες έχουν εκτιμηθεί με χρήση συντελεστή κάλυψης ίσο με 2 που αντιστοιχεί σε επίπεδο εμπιστοσύνης 95% περίπου.

6. Συμπεράσματα

Στα πλαίσια της παρούσας εργασίας έγινε εφαρμογή και σύγκριση τριών εναλλακτικών στατιστικών προσεγγίσεων για την ανάλυση δεδομένων μετρήσεων με στόχο την εκτίμηση της αβεβαιότητας δειγματοληψίας καυσίμων. Υλοποιήθηκε ένα ισορροπημένο πειραματικό σχέδιο, το οποίο περιελάμβανε λήψη διπλών δειγμάτων ντίζελ κίνησης από 104 πρατήρια καυσίμων (10,9% των πρατηρίων που παρακολουθούνται από το εργαστήριο) και στη συνέχεια εις διπλούν αναλύσεις αυτών των δειγμάτων για τον προσδιορισμό της περιεκτικότητάς τους σε θείο.

Τα αποτελέσματα των μετρήσεων επεξεργάστηκαν με τη χρήση «κλασικής» ANOVA, ανθεκτικής ANOVA και Στατιστικής Εύρους Τιμών. Οι τρεις μεθοδολογίες έδωσαν διαφορετικές εκτιμήσεις με τη διευρυμένη αβεβαιότητα της δειγματοληψίας να κυμαίνεται από 0,34 έως 0,40 mg kg⁻¹, ενώ η σχετική διευρυμένη αβεβαιότητα από 4,8 έως 5,1 % . Το γεγονός ότι η ανθεκτική ANOVA οδηγεί σε διαφορετικά αποτελέσματα σε σύγκριση με τις άλλες δύο μεθοδολογίες είναι μια ένδειξη ότι οι παραδοχές για την εφαρμογή της «κλασικής» ANOVA και της Στατιστικής Εύρους Τιμών δεν ισχύουν. Το γεγονός αυτό επιβεβαιώνεται και από την παρουσία ενός μικρού αριθμού ακραίων τιμών (8,7 %) εντός των δεδομένων που χρησιμοποιήθηκαν για τους υπολογισμούς. Ως εκ τούτου, η ανθεκτική ANOVA, η οποία δεν επηρεάζεται από ακραίες τιμές, θεωρείται ως η μέθοδος που παρέχει τις πιο αξιόπιστες εκτιμήσεις για τη διευρυμένη αβεβαιότητα δειγματοληψίας (0,34 mg kg⁻¹ ή 4,8 % ως σχετική διευρυμένη αβεβαιότητα). Τα αποτελέσματα της ανθεκτικής ANOVA δείχνουν ότι η αβεβαιότητα λόγω ανάλυσης αντιπροσωπεύει το 71 %, ενώ η αβεβαιότητα λόγω δειγματοληψίας αντιπροσωπεύει μόνο το 29 % της συνολικής αβεβαιότητας της μέτρησης.

ΒΙΒΛΙΟΓΡΑΦΙΑ

ASTM Standard D 4057-12, “*Standard Practice for Manual Sampling of Petroleum and Petroleum Products*”, ASTM International, West Conshohocken, PA, 2012a.

ASTM Standard D 5453-12, “*Standard Test Method for Determination of Total Sulfur in Light Hydrocarbons, Spark Ignition Engine Fuel, Diesel Engine Fuel, and Engine Oil by Ultraviolet Fluorescence*”, ASTM International, West Conshohocken, PA, 2012b.

Ellison S.L.R., Barwick V.J., Farrant T.J.D., “*Practical Statistics for the Analytical Scientist: A Bench Guide*”, second ed., The Royal Society of Chemistry, Cambridge, UK, 2010.

Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide, “*Measurement Uncertainty arising from sampling: a guide to methods and approaches*”, 2007 (Available from the Eurachem secretariat at <http://www.eurachem.org>)

Heydorn K., Esbensen K., “*Sampling and metrology*”, Accreditation and Quality Assurance, 9, p.p. 391-396, 2004.

ISO 20846, “*Petroleum products - Determination of sulfur content of automotive fuels - Ultraviolet fluorescence method International Organization for Standardization*”, Geneva, Switzerland, 2011.

ISO 3085, “*Iron ores - Experimental methods for checking the precision of sampling, sample preparation and measurement*”, International Organization for Standardization, Geneva, Switzerland, 2002.

Lyn J.A., Ramsey M.H., Damant A.P., Wood R., “*Empirical versus modelling approaches to the estimation of measurement uncertainty caused by primary sampling*”, Analyst, 132, p.p.1231-1237, 2007.

Nordtest, NT Technical Report 604, “*Uncertainty from sampling. A Nordtest handbook for sampling planners on sampling quality assurance and uncertainty estimation*”, Nordisk Innovation Centre, Oslo, Norway, 2007.

Ramsey M. H., “*Objective evaluation of precision requirements for geochemical analysis using robust analysis of variance*”, Journal of Geochemical Exploration, 44, p.p. 23-36, 1992.

Ramsey M. H., “*Sampling as a source of measurement uncertainty: Techniques for quantification and comparison with analytical sources*”, Journal of Analytical Atomic Spectrometry, 13, p.p. 97-104, 1998.

Ramsey M.H., “*When is sampling part of the measurement process?*”, Accreditation and Quality Assurance, 9, p.p. 727-728, 2004.

Royal Society of Chemistry – Analytical Methods Committee, ROBAN, <http://www.rsc.org/Membership/Networking/InterestGroups/Analytical/AMC/Software/ROBAN.asp> (Accessed Jan 3, 2014)

Royal Society of Chemistry – Analytical Methods Committee, “*Robust statistics—how not to reject outliers. Part 1. Basic concepts*”, Analyst, 114, p.p.1693-1697, 1989.