Η ΕΚΤΙΜΗΣΗ ΤΗΣ ΜΕΤΡΗΤΙΚΗΣ ΑΒΕΒΑΙΟΤΗΤΑΣ ΣΕ ΠΕΡΙΠΤΩΣΕΙΣ ΥΠΑΡΞΗΣ ΣΥΣΧΕΤΙΣΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ

Χ. Λ. Μήτσας

Διεύθυνση Μηχανικών Μεγεθών, Ε.Ι.Μ., Θεσσαλονίκη, 57022
E-mail: chris.mitsas@eim.gr

Περίληψη
Στην παρούσα εργασία εξετάζεται η επίδραση της συν-διακύμανσης μεταξύ συσχετισμένων μετρήσεων στην αβεβαιότητα του αποτελέσματος μέτρησης καθώς και αποτελεσματικοί τρόποι αντιμετώπισης μερικών συχνά εμφανιζόμενων περιπτώσεων συσχέτισης. Το πλαίσιο σε αυτή την απόπειρα παρέχεται από το ISO “GUM”. Επίσης, παρουσιάζεται ένα παράδειγμα δειγματοληψίας συσχετισμένων τυχαίων μεταβλητών για εφαρμογή σε προσομοιώσεις και επεκτείνεται η χρήση υπολογιστικών φύλλων για την εκτίμηση αβεβαιότητας σε περιπτώσεις ύπαρξης συσχετίσεων μεταξύ των μετρούμενων μεγεθών.

Abstract
The present work investigates the influence of covariance between correlated measurements on the estimation of the uncertainty of a measurement result as well as efficient methods for handling the most common types of correlation. The ISO “GUM” provides the framework for this study. In addition, examples are presented concerning the generation of samples of correlated random variables for use in simulation as well as for a spreadsheet method of uncertainty budgeting including covariance between the model variables.

1. Εισαγωγή
Το “Guide to the Expression of Uncertainty in Measurements” (GUM) (ISO 1995) αποτελεί εδώ και 15 χρόνια την διεθνώς αποδεκτή μεθοδολογία για την εκτίμηση της αβεβαιότητας μετρήσεων, μιάς παραμέτρου που χαρακτηρίζει την αξιοπιστία του αποτελέσματος μέτρησης ενός φυσικού μεγέθους. Αν και η μεθοδολογία πλέον εφαρμόζεται στην ευρύτερη μετρολογική κοινότητα, σε κάποιες περιπτώσεις η ορθότητα της εφαρμογής μπορεί να θέσει υπό αμφισβήτηση το τελικό αποτέλεμα της μέτρησης. Η συνεκτίμηση της επίδρασης συσχετίσεων μεταξύ των μετρούμενων μεγεθών ή αντίστοιχα συν-διακυμάνσεων των τυχαίων μεταβλητών που υπεισέρχονται στο μοντέλο που περιγράφει την μετρητική διεργασία είναι μία τέτοια περίπτωση. Συνήθως υπάρχει απροθυμία συμπερίληψης αυτού του είδους «επιπλοκών» στο ισοζύγιο αβεβαιότητας θεωρώντας είτε “de facto” την απουσία συσχετισμών ή απλώς αγνοώντας την συνεισφορά τους.

Στην πραγματικότητα, σπάνια εμφανίζονται μετρήσεις χωρίς συσχετίσεις, αν και η σημαντικότητά τους μπορεί να διαφέρει κατά περίπτωση. Αυτές οφείλονται πάντα σε συστηματικές επιδράσεις, είτε άμμεσες μεταξύ των μετρούμενων μεγεθών ή έμμεσες μέσω ενός τρίτου μεγέθους. Έτσι, οι συσχετίσεις μεταξύ μετρούμενων μεγεθών μπορούν να αποδοθούν είτε στην ταυτόχρονη μέτρησή τους είτε στην υποβόσκουσα επίδραση από την χρήση κοινού μετρητικού εξοπλισμού, προτύπου ή δεδομένου αναφοράς, μετρητικής μεθόδου ή ακόμη και παρόχου υπηρεσιών διακρίβωσης.

Στην παρούσα εργασία, γίνεται μία σύστηματική μελέτη της μεθοδολογίας του “GUM” σε περιπτώσεις ύπαρξης συσχετισμένων μετρήσεων. Ιδιαίτερα εξετάζονται περιπτώσεις με συχνή εμφάνιση στον χώρο των διακριβώσεων φυσικών μεγεθών με σκοπό την ανάδειξη καταστάσεων όπου η συσχέτιση μπορεί να επιδράσει παρατηρήσιμα στην αξιοπισία του αποτελέσματος. Συμπληρωματικά, παρουσιάζεται ένα παράδειγμα παραγωγής τυχαίων δειγμάτων από μία κανονική δι-μεταβλητή κατανομή για χρήση στην διάδοση κατανομών πιθανοτήτων με προσομοίωση Monte-Carlo (JCGM 2008) καθώς και μία τροποποίηση της μεθόδου εκτίμησης αβεβαιότητας μέσω υπολογιστικού φύλλου (Kragten 1994) ώστε να περιλαμβάνει συνδιακυμάνσεις μεταξύ των μεταβλητών του μοντέλου μέτρησης.

2. Η Μεθοδολογία του Καθιερωμένου GUM για Συσχετισμένες Μετρήσεις

Ένα από τα μεγαλύτερα πλεονεκτήματα της μεθοδολογίας του "GUM" είναι το γεγονός ότι η εκτίμηση της μετρητικής αβεβαιότητας προέρχεται από το ίδιο μοντέλο που χρησιμοποιείται για τον υπολογισμό του αποτελέσματος της μέτρησης. Δεδομένης της ευελιξίας στην επιλογή του μοντέλου καθώς και της εισαγωγής σε αυτό όλης της διαθέσιμης πληροφορίας σχετικά με την μέτρηση, η μεθοδολογία μπορεί να χρησιμοποιηθεί για την επιτυχή αντιμετώπιση σχεδόν κάθε περίπτωσης μέτρησης. Η διαθέσιμη γνώση, αφού μετατραπεί σε τυπική αβεβαιότητα $u(x_i)$ ή συν-διακύμανση, $u(x_i, x_j)$ των εκτιμήσεων x_i, x_j, διαδίδεται μέσω του νόμου διάδοσης των αβεβαιοτήτων για να εκτιμηθεί η τυπική αβεβαιότητα του αποτελέσματος, $u(y)$ μέσω της έκφρασης (παρ. 5.2 του "GUM")

$$u_c^2(y) = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}u(x_i, x_j) \tag{1}$$

Η παραπάνω σχέση μπορεί να γραφεί ισοδύναμα με μορφή πινάκων

$$u(y) = \mathbf{J}\mathbf{u}(x_i, x_j)\mathbf{J}^T \tag{2}$$

όπου **J** είναι ο επονομαζόμενος Ιακωβιανός πίνακας και ισούται με $\left(\frac{\partial f}{\partial x_1}\cdots\frac{\partial f}{\partial x_N}\right)$ ενώ με T συμβολίζεται ο ανάστροφος. Αυτή η μορφή έχει σαφή πλεονεκτήματα ιδιαίτερα σε υπολογιστικές εφαρμογές ενώ εύκολα επεκτείνεται σε περιπτώσεις πολυμεταβλητών αποτελεσμάτων μέτρησης.

Η συν-διακύμανση πρακτικά εκφράζεται μέσω του συντελεστή συσχέτισης $r(x_i, x_j)$ ο οποίος παίρνει τιμές μεταξύ $-1 \le r(x_i, x_j) = r(x_j, x_i) \le 1$ ή του αντίστοιχου NxN πίνακα $\mathbf{r(x_i, x_j)}$

$$r(x_i, x_j) = \frac{u(x_i, x_j)}{u(x_i)\,u(x_j)} \qquad \text{ή} \qquad \mathbf{u}(x_i, x_j) = \mathbf{u}(x_i)\mathbf{r}(x_i, x_j)\mathbf{u}(x_j)^T \tag{3}$$

περιγράφοντας με έναν περιεκτικό τρόπο τον βαθμό συσχέτισης μεταξύ των εκτιμήσεων των μεταβλητών του μοντέλου μέτρησης. Το βασικό ζητούμενο για πληρότητα του παραπάνω φορμαλισμού είναι τρόποι εκτίμησης των συντελεστών συσχέτισης. Αν και δεν αποδίδεται πάντα η δέουσα προσοχή, το "GUM" περιλαμβάνει σχετικές τεχνικές ανάλογα με τον τρόπο υπολογισμού των αβεβαιοτήτων.

Έτσι στην περίπτωση επαναληπτικών μετρήσεων (παρ. 5.3.2), που συνάδει με τον τρόπο Α υπολογισμού αβεβαιοτήτων, αναφέρει ότι η συν-διακύμανση μεταξύ συσχετισμένων μεταβλητών X_i, X_j στο μοντέλο μέτρησης που εκτιμούνται από τους μέσους όρους n ανεξάρτητων ζευγών ταυτόχρονων παρατηρήσεων κάτω από τις ίδιες συνθήκες μέτρησης υπολογίζεται ως

$$u(x_i, x_j) = s(\overline{X}_i, \overline{X}_j) \approx \frac{1}{n(n-1)}\sum_{k=1}^{n}(x_{ik} - \overline{X}_i)(x_{jk} - \overline{X}_j) \tag{4}$$

Αντίστοιχα, η περίπτωση όπου η συσχέτιση προέρχεται έμμεσα από ένα τρίτο μέγεθος και η αβεβαιότητα εκτιμάται με τον τρόπο Β υπολογισμού εκτίθεται στις παρ. 5.2.4 και F.1.2.3. Αν λοιπόν δύο μεταβλητές X_i, X_j εξαρτώνται από έναν αριθμό ασυσχέτιστων μεγεθών $Q_1,\ldots, Q_L$ μέσω γνωστών σχέσεων $X_i = F(Q_1,\ldots, Q_L)$ και $X_j = G(Q_1,\ldots, Q_L)$, χωρίς υποχρεωτικά να εμφανίζονται όλες οι μεταβλητές Q_l και στις δύο παραπάνω συναρτησιακές σχέσεις, τότε οι αβεβαιότητες των εκτιμήσεών τους είναι

$$u^2(x_i) = \sum_{l=1}^{L}\left(\frac{\partial F}{\partial q_l}\right)^2 u^2(q_l) \qquad\qquad u^2(x_j) = \sum_{l=1}^{L}\left(\frac{\partial G}{\partial q_l}\right)^2 u^2(q_l)$$

όπου $u(q_l)$ η αβεβαιότητα της εκτίμησης της μεταβλητής Q_l. Εφαρμόζοντας την διάδοση αβεβαιοτήτων στο μοντέλο που τώρα περιλαμβάνει μόνο ασυσχέτιστες μεταβλητές $Y = f(X_1,\ldots, X_{i-1}, X_{i+1}, X_{j-1}, X_{j+1},\ldots, X_N, Q_1,\ldots, Q_L)$, προκύπτει ότι

$$u(x_i, x_j) = \sum_{l=1}^{L} \frac{\partial F}{\partial q_l} \frac{\partial G}{\partial q_l} u^2(q_l) \qquad (5)$$

Πρέπει να σημειωθεί η παρατήρηση στις παραγράφους 5.2.4 και F.1.2.4 του "GUM" ότι, όπου είναι εφικτό, πρέπει να προτιμάται να περιλαμβάνονται αναλυτικά οι κοινές επιδράσεις ως ανεξάρτητες μεταβλητές στο μετρητικό μοντέλο με αποτέλεσμα την αποφυγή συσχετίσεων και των αντίστοιχων συν-διακυμάνσεων στην εκτίμηση της αβεβαιότητας του αποτελέσματος.

3. Μερικά Παραδείγματα Εφαρμογής

3.1 Προσθετικά και Παραγοντικά Μοντέλα

Πολύ συχνά η διεργασία της διακρίβωσης μπορεί να παρασταθεί με απλά προσθετικά ή παραγοντικά μοντέλα μέτρησης της μορφής

$$Y = c_1 X_1 \pm c_2 X_2 \qquad \text{ή} \qquad Y = C X_1^{p_1} X_2^{p_2}$$

Το αθροιστικό μοντέλο με το πρόσημο «+» θα μπορούσε να αντιπροσωπεύει για παράδειγμα την ταυτόχρονη χρήση δύο προτύπων βαρών για την διακρίβωση της ένδειξης ζυγού ενώ με το πρόσημο «-» θα μπορούσε να αντιπροσωπεύει μία σύγκριση προτύπων βαρών με ίδιες ονομαστικές τιμές.

Εφαρμόζοντας την (1) για την εκτίμηση της αβεβαιότητας του αποτελέσματος προκύπτουν εντελώς αντίστοιχες εκφράσεις

$$u(y) = \sqrt{c_1^2 u^2(x_1) + c_2^2 u^2(x_2) \pm 2 c_1 c_2 u(x_1) u(x_2) r(x_1, x_2)} \qquad (6\alpha)$$

$$\frac{u(y)}{y} = \sqrt{p_1^2 \xi^2 + p_2^2 \eta^2 \pm 2 p_1 p_2 (\xi)(\eta) r(\xi, \eta)} \qquad (6\beta)$$

όπου οι μεταβλητές ξ και η στην (6β) εκφράζουν τις αντίστοιχες σχετικές αβεβαιότητες των εκτιμήσεων x_1 και x_2. Αν οι μεταβλητές είναι πλήρως συσχετισμένες, $r = \pm 1$, τότε η αβεβαιότητα των αποτελεσμάτων δίνεται από τις απλές εκφράσεις

$$u(y) = \left|(c_1 u(x_1) \pm c_2 u(x_2)\right| \qquad \text{και} \qquad \frac{u(y)}{y} = \left|p_1 \xi \pm p_2 \eta\right|$$

Είναι προφανές ότι ανάλογα με τα πρόσημα είτε των συντελεστών c_1, c_2 ή των εκθετών p_1, p_2 η αβεβαιότητα μπορεί να είναι σημαντικά διαφορετική από αυτή που εκτιμάται στην περίπτωση ασυσχέτιστων μεταβλητών. Η εξάρτηση της διακύμανσης V(y) από τον συντελεστή συσχέτισης r και από τον λόγο των επιμέρους αβεβαιοτήτων των μεταβλητών $V_{rel} = u(x_1)/u(x_2)$ του αθροιστικού μοντέλου με $c_1 = c_2 = 1$ φαίνεται στο σχήμα 1 (σημεία). Επιπλέον, απεικονίζεται η αντίστοιχη διακύμανση με την υπόθεση της ανεξαρτησίας των μεταβλητών (γραμμές). Συμπερασματικά, η συμπερίληψη της συν-διακύμανσης στο αποτέλεσμα της μέτρησης είναι σημαντική όταν οι αβεβαιότητες των μεταβλητών του μοντέλου είναι παρόμοιες ($V_{rel} \leq 3$), ενώ σε αντίθετα περίπτωση δεν επιδρά σημαντικά στην εκτίμηση.

3.2 Ταυτόχρονος Προσδιορισμός Αριθμού Μετρούμενων Μεγεθών

Στο ISO "GUM" αναφέρεται ότι μία εκτίμηση ενός μετρούμενου μεγέθους Y μέσω του μετρητικού μοντέλου μπορεί να γίνει με δύο ισοδύναμους τρόπους (παρ. 4.1.4)

α) από τους μέσους όρους των N μεταβλητών που προκύπτουν από n ανεξάρτητες παρατηρήσεις έκαστης εξ' αυτών

$$y = f(\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_N)$$

β) από τον μέσο όρο n ανεξάρτητων προσδιορισμών Y_k, του Y, που έχουν την ίδια αβεβαιότητα και προκύπτουν από την ταυτόχρονη παρατήρηση των N μεταβλητών

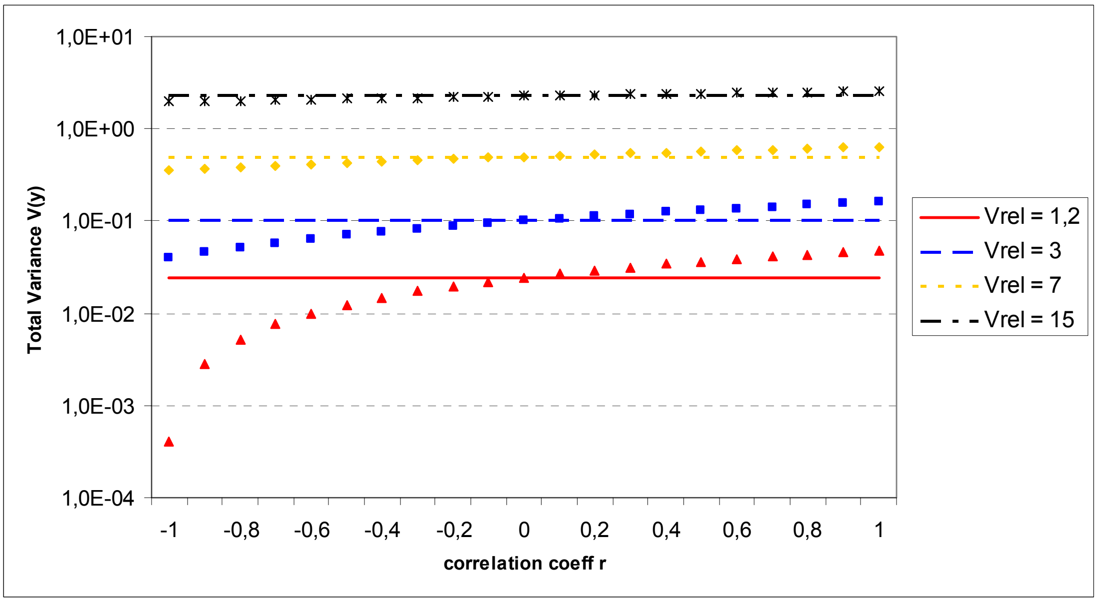
$$y = \overline{Y} = \frac{1}{n}\sum_{k=1}^{n} f(X_{1,k}, X_{2,k}, \ldots, X_{N,k})$$

Το "GUM" προτείνει την χρήση του (β) τρόπου όταν το μοντέλο είναι μη-γραμμικό ή ιδιαίτερα πολύπλοκο ενώ στα Παραρτήματα H.2 και H.4 του ιδίου αναπτύσσονται παραδείγματα εκτίμησης της αβεβαιότητας καθώς και των συν-διακυμάνσεων και με τους δύο τρόπους.

Το παράδειγμα που θα παρουσιασθεί εδώ αφορά στον προσδιορισμό της πυκνότητας του αέρα, ρ_a, η οποία χρησιμοποιείται για τον υπολογισμό της διόρθωσης άνωσης στον προσδιορισμό μάζας. Σε περιπτώσεις υψηλών απαιτήσεων, η πυκνότητα του αέρα υπολογίζεται μέσω της αναθεωρημένης σχέσης CIPM 81/91 (Davis 1992, Park et. al. 2004), αλλά για τους σκοπούς του παραδείγματος θα εκτιμηθεί από μία προσεγγιστική σχέση (Schwartz et. al. 2007)

$$\rho_a = \frac{0{,}34848 \cdot p - 0{,}009024 \cdot h \cdot \exp(0{,}0612t)}{273{,}15 + t} \qquad (7)$$

Στον πίνακα 1 παρατίθενται οι ταυτόχρονες μετρήσεις της θερμοκρασίας του αέρα, t, της σχετικής υγρασίας, h, και της ατμοσφαιρικής πίεσης, p για τον προσδιορισμό της πυκνότητας του αέρα μέσω της (7). Στον ίδιο πίνακα παρουσιάζονται οι εκτιμήσεις του μετρούμενου μεγέθους και οι αντίστοιχες αβεβαιότητες σύμφωνα με τους τρόπους που προτείνονται από το "GUM": τρόπος (α), $\rho_a(\alpha)$ και $u(\rho_a)$, και τρόπος (β), $\rho_a(\beta)$ και $s_m(\rho_a)$. Επίσης, με σκοπό την σύγκριση, φαίνεται και η αβεβαιότητα, $u_{uncorr}(\rho_a)$, υπολογισμένη θεωρώντας ότι δεν υπάρχουν συσχετίσεις μεταξύ των μεταβλητών του μοντέλου. Από την σύγκριση προκύπτει ότι και οι δύο τρόποι παράγουν ισοδύναμα αποτελέσματα ήτοι, 1,18770 ± 4,9x10^{-4} kg/m^3. Αντίθετα, η παράλειψη της συν-διακύμανσης των μεταβλητών του μοντέλου υποεκτιμά την αβεβαιότητα περίπου 10% σε αυτή τη περίπτωση. Ο πίνακας των συντελεστών συσχέτισης, $r(x_i, x_j)$, που χρησιμοποι υπολογίστηκε μέσω των (3) και (4).



Σχήμα 1. Εξάρτηση της διακύμανσης V(y) από τον συντελεστή συσχέτισης r και από τον λόγο V_{rel} . Συσχετισμένες μεταβλητές (σημεία), ανεξάρτητες μεταβλητές (γραμμή)

Πίνακας 1. Μετρήσεις θερμοκρασίας αέρα, t, σχετικής υγρασίας, h, και ατμοσφαιρικής πίεσης, p και συγκριτική εκτίμηση μετρητικής αβεβαιότητας της πυκνότητας του αέρα.

Θερμοκρασία Αέρα t (^oC)	Σχετική Υγρασία (%)	Ατμοσφαιρική Πίεση (mbar)	Πυκνότητα Αέρα $\rho_{a,i}$ (kg/m^3)
23,3	45,5	1015,1	1,18750
23,2	47,1	1015,3	1,18797
23,1	49,3	1015,9	1,18883
22,9	51,0	1015,5	1,18903
23,2	48,6	1014,7	1,18707
23,3	47,7	1013,9	1,18581

$\overline{X}_i$ =	23,17	48,20	1015,07	$\rho_a(\beta)$ =	1,18770
$s_m(X_i)$ =	0,06	0,77	0,28	$s_m(\rho_a)$ =	4,9E-04
$\rho_a(\alpha)$ =	1,18770				
$u(\rho_a)$ =	4,9E-04				
$u_{uncorr}(\rho_a)$ =	4,4E-04				

$$r(x_i, x_j) = \begin{vmatrix} 1{,}000 & -0{,}890 & -0{,}622 \\ -0{,}890 & 1{,}000 & 0{,}336 \\ -0{,}622 & 0{,}336 & 1{,}000 \end{vmatrix}$$

3.3 Αναλυτικά Μοντέλα με Γνωστές Συσχετίσεις

Όταν οι μεταβλητές του μοντέλου σχετίζονται με τύπου Β υπολογισμό αβεβαιοτήτων και υπάρχει κάποιο επικουρικό μοντέλο το οποίο περιγράφει τη συσχέτιση μεταξύ τους τότε μπορεί να χρησιμοποιηθεί η σχέση (6) για την εκτίμηση της συνδιακύμανσής τους. Σε αυτή την περίπτωση είναι προφανές ότι θα υπάρχουν συν-διακυμάνσεις μόνο μέσω κοινών μεταβλητών σε όλες τις συναρτησιακές σχέσεις του βοηθητικού μοντέλου. Για την επίδειξη της μεεθόδου θα χρησιμοποιηθεί ένα παράδειγμα από την μετρολογία μάζας. Ως γνωστό, η μάζα ενός αγνώστου πρότυπου βάρους m_t, προσδιορίζεται σε σχέση με την μάζα ένα πρότυπο βάρος αναφοράς m_r, ίδιας ονομαστικής τιμής μέσω ζύγισης όπως περιγράφεται από την σχέση

$$m_t = m_r + \rho_a (V_t - V_r) + \Delta m_w \tag{8}$$

όπου V_t και V_r είναι οι όγκοι του αγνώστου βάρους και του προτύπου βάρους αναφοράς αντίστοιχα και Δm_w η διαφορά ένδειξης του ζυγού αφού μετατραπεί σε μάζα μέσω της ευαισθησίας του (Schwartz et. al. 2007). Το πρότυπο αναφοράς συχνά παίζει τον ρόλο ενός προτύπου μεταφοράς της ακρίβειας αφού και αυτό έχει διακριβωθεί σε σχέση με άλλο πρότυπο μάζας m_{st} και όγκου V_{st} υψηλότερης ακρίβειας σε πυκνότητα αέρα ρ_{st} ώστε

$$m_r = m_{st} + \rho_{st} (V_r - V_{st}) + \Delta m_{w,st}$$

Το παραπάνω υποδηλώνει ότι η μάζα και ο όγκος του προτύπου μεταφοράς, m_r και V_r, είναι συσχετισμένα μεγέθη και ως εκ τούτου χαρακτηρίζονται από συν-διακύμανση που δίνεται από

$$u(m_r, V_r) = \frac{\partial m_r}{\partial V_r} \frac{\partial V_r}{\partial V_r} u^2(V_r) = \rho_{st} u^2(V_r)$$

ενώ ο συντελεστής συσχέτισης μέσω της (3) γράφεται ως

$$r(m_r, V_r) = \rho_{st} \frac{u(V_r)}{u(m_r)} \tag{9}$$

Για αντιπροσωπευτικές τιμές του εργαστηρίου Μάζας του ΕΙΜ, ο συντελεστής της (9) ισούται με 0,2207. Με βάση τα παραπάνω, η αβεβαιότητα της διόρθωσης άνωσης u_{buoy}, συμπεριλαμβανομένης της συνδιακύμανσης $u(m_r, V_r)$ δίνεται από

$$u_{buoy}^2 = (V_t - V_r)\,u^2(\rho_a) + \rho_a^2 u^2(V_x) + (\rho_a^2 - 2\rho_a\rho_{st})\,u^2(V_r)$$

Είναι χαρακτηριστικό ότι στην περίπτωση που $\rho_a = \rho_{st}$ αφαιρείται η συνεισφορά της αβεβαιότητας του όγκου του προτύπου μεταφοράς δεδομένου ότι η άνωση που αυτό υφίσταται δεν συμμετέχει καθόλου στον προσδιορισμό της μάζας του αγνώστου πρότυπου βάρους.

3.4 Προσαργογή Μέσω του Κριτηρίου των Ελαχίστων Τετραγώνων

Ο εκτιμητής των ελαχίστων τετραγώνων είναι ίσως ο δημοφιλέστερος εκτιμητής στηριζόμενος στην αρχή του πιθανότερου ενδεχομένου (maximum likelihood estimator) (Taylor 1992) και ως εκ τούτου δεν θα μπορούσε να απουσιάζει από το "GUM" (Παράρτημα H.3). Επειδή υπάρχει εκτενής βιβλιογραφία όσον αφορά στην εκτίμηση μέσω ελαχίστων τετραγώνων (Mandel 1984, Barlow 1989 για παράδειγμα) δεν θα επιχειρηθεί συνολική ανασκόπηση της μεθόδου, αλλά θα εστιάσουμε στην αξιοπιστία της πρόβλεψης που γίνεται μέσω των εκτιμηθέντων παραμέτρων ενός απλού γραμμικού μοντέλου μεταξύ δύο μεταβλητών Y και X.

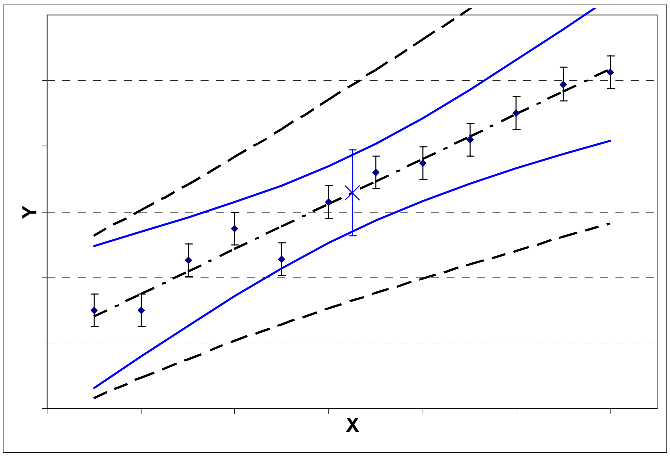
Ξεκινώντας από τις στατιστικές υποθέσεις του προβλήματος, δηλ., ότι τα σφάλματα των παρατηρήσεων y_i είναι ανεξάρτητα και ομοιόμορφα κατανεμημένα και προέρχονται από μία κανονική κατανομή *N(0, σ)*, η ευθεία που περιγράφει βέλτιστα την σχέση μεταξύ των δύο μεταβλητών είναι

$$\hat{\mathbf{Y}} = \hat{m}X + \hat{c} \qquad (10)$$

όπου $\hat{m}$, $\hat{c}$ οι εκτιμήσεις των παραμέτρων της και $\hat{Y}$ η πρόβλεψη του μοντέλου για τιμές της μεταβλητής X. Η αβεβαιότητα της πρόβλεψης μπορεί απλά να υπολογιστεί από την εφαργογή της διάδοσης των αβεβαιοτήτων στην (21), λαμβάνοντας υπόψη όμως ότι οι εκτιμηθείσες παράμετροι είναι συσχετισμένες μέσω των παρατηρήσεων y_i, ώστε αυτή να δίνεται ως

$$u(\hat{Y}) = \sqrt{X^2 u^2(\hat{m}) + u^2(\hat{c}) + 2Xu(\hat{m},\hat{c})}$$

Ο τελευταίος όρος της συν-διακύμανσης στην παραπάνω σχέση είναι πολύ σημαντικός και εάν δεν συμπεριληφθεί θα οδηγήσει σε υπερεκτίμηση της αβεβαιότητας της πρόβλεψης. Στο σχήμα 2. απεικονίζονται τα αποτελέσματα προσαρμογής δεδομένων με χρήση των ελαχίστων τεραγώνων, καθώς επίσης και η ζώνη εμπιστοσύνης (95%) που αποδίδεται στις προβλέψεις του γραμμικού μοντέλου όταν περιλαμβάνεται ή όχι η συνδιακύμανση.



Σχήμα 2. Ελάχιστα τετράγωνα σε γραμμικό μοντέλο. Οι καμπύλες απεικονίζουν ζώνες επμιστοσύνης (95%) με (συμπαγής) ή χωρίς (διακεκομένη) την συνδιακύμανση. Με τον σταυρό απεικονίζεται το κέντρο βάρους των ζευγών και η αντίστοιχη αβεβαιότητα (95%).

Αποδεικνύεται (Draper and Smith 1998) ότι αν τα ζεύγη (x_i, y_i) κεντρωθούν, δηλ. η προσαρμογή γίνει στα $(x_i - \overline{X}, y_i - \overline{Y})$, όπου $(\overline{X}, \overline{Y})$ το κέντρο βάρους των πειραματικών ζευγών, τότε οι προβλέψεις του μοντέλου και οι αντίστοιχες διακυμάνσεις δίνονται από

$$\hat{Y} = \hat{m}(X - \overline{X}) + \overline{Y}$$

$$u^2(\hat{Y}) = (X - \overline{X})^2 u^2(\hat{m}) + u^2(\overline{Y})$$

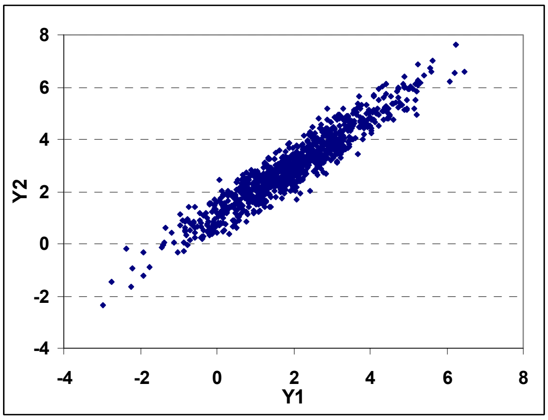
αφού $u(\overline{Y}, \hat{m}) = 0$. Από την τελευταία σχέση συμπεραίνουμε ότι για $X = \overline{X}$ η αβεβαιότητα της πρόβλεψης ισούται με τη τυπική απόκλιση του μέσου $\overline{Y}$.

4. Προσομοιώση Monte-Carlo με Συσχετισμένες Μεταβλητές

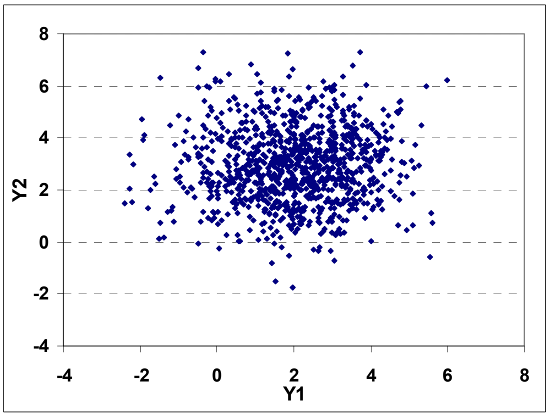
Η διάδοση των κατανομών των τυχαίων μεταβλητών ενός μετρητικού μοντέλου μέσω προσομοίωσης για την εκτίμηση της αβεβαιότητας και του αντίστοιχου διαστήματος εμπιστοσύνης του αποτελέσματος της μέτρησης περιγράφεται στο “GUM S1” (JCGM 2008). Όταν οι τυχαίες μεταβλητές είναι συσχετισμένες η δειγματοληψία τυχαίων αριθμών γινεται από την κοινή συνάρτηση πυκνότητας πιθανοτήτων (joint PDF) της αντίστοιχης πολυμεταβλητής κατανομής. Το “GUM S1” περιγράφει ένα σχετικό αλγόριθμο γεννήτριας τυχαίων αριθμών για την πολυμεταβλητή κανονική κατανομή, $N(\boldsymbol{\mu}, V)$ όπου **μ** είναι ένα nx1 διάνυσμα αναμενόμενων τιμών και **V** ένας nxn πίνακας συν-διακύμανσης (παρ. C.5). Προϋπόθεση για την λειτουργία της γεννήτριας είναι ο πίνακας **V** να είναι τουλάχιστο θετικά ημι-ορισμένος, ώστε ο αλγόριθμος να μην επηρεάζεται από διάφορες πηγές αριθμητικών σφαλμάτων (Kessel et. al. 2009). Στο σχήμα 3α παρουσιάζεται η δειγματοληψία 1000 σημείων από μία δι-μεταβλητή (bi-variate) κανονική κατανομή με θετική συσχέτιση, αντίστοιχη του σχήματος C.1 που εμφανίζεται στο “GUM S1”. Οι αναμενόμενες τιμές των μεταβλητών, **μ**, και ο πίνακας συν-διακύμανσης, **V**, που χρησιμοποιήθηκαν ήταν

$$\boldsymbol{\mu} = \begin{pmatrix} 2{,}0 \\ 3{,}0 \end{pmatrix} \qquad \mathbf{V} = \begin{pmatrix} 2{,}0 & 1{,}9 \\ 1{,}9 & 2{,}0 \end{pmatrix}$$

και η προσομοίωση πραγματοποιήθηκε με το πρόσθετο του EXCEL MCSim (Barreto and Howland 2005). Η θετική κλίση είναι απόρροια της θετικής συσχέτισης (r = 0,95) μεταξύ των μεταβλητών. Στο σχήμα 3β φαίνεται η ίδια προσομοίωση με ανεξάρτητες μεταβλητής (r = 0).



Σχήμα 3α. Ζεύγη (Y_1, Y_2) που λήφθησαν από δι-μεταβλητή κανονική κατανομή (r = 0,95).



Σχήμα 3b. Ζεύγη (Y_1, Y_2) που λήφθησαν από ανεξάρτητες κανονικές κατανομές (r = 0).

5. Συσχετισμένες Μετρήσεις και Υπολογιστικά Φύλλα

Η ιδέα της χρήσης υπολογιστικών φύλλων (spreadsheets) για την κατάρτιση ισοζυγίων αβεβαιότητας και αριθμητικής εφαρμογής του νόμου διάδοσης των αβεβαιοτήτων προτάθηκε πρίν από περίπου 15 χρόνια (Kragten 1994). Το σκεπτικό είναι ότι αν η αβεβαιότητα $u(x_i)$ θεωρηθεί ως μία μικρή μεταβολή της εκτίμησης x_i, τότε μέσω του αναπτύγματος Taylor πρώτης τάξης της συνάρτησης του μετρητικού μοντέλου προκύπτει ότι

$$\Delta f_i = f[x_i + u(x_i)] - f(x_i) \approx \frac{\partial f}{\partial X_i} u(x_i) \qquad (11)$$

Αμέσως γίνεται αντιληπτό ότι οι όροι που υπεισέρχονται στον νόμο διάδοσης αβεβαιοτήτων μπορούν να προσεγγισθούν, κάτω από προϋποθέσεις, από μία απλή διαφορά των τιμών της συνάρτησης αποφεύγοντας περίπλοκους αναλυτικούς υπολογισμούς ή επικίνδυνα αριθμητικά σφάλματα. Επιπλέον, η (2) με την βοήθεια της (3) και (11) γράφεται ως

$$u(y) = (\Delta f_1 \dots \Delta f_N) \begin{pmatrix} r_{11} & \cdots & r_{1N} \\ \vdots & \vdots & \vdots \\ r_{N1} & \cdots & r_{NN} \end{pmatrix} \begin{pmatrix} \Delta f_1 \\ \vdots \\ \Delta f_N \end{pmatrix}$$

Με την προϋπόθεση ότι ο πίνακας συσχέτισης είναι γνωστός η αβεβαιότητα του αποτελέσματος μπορεί να υπολογισθεί εύκολα μέσω πολλαπλασιαμού πινάκων στο EXCEL (Gottfried 1998).

Συμπεράσματα

Παρουσιάσθηκαν οι κυριότερες πηγές συν-διακύμανσης συσχετισμένων μετρήσεων και διάφορες τεχνικές για τον συν-υπολογισμό της στην αβεβαιότητα του μετρητικού αποτελέσματος. Οι περισσότερες από αυτές τις τεχνικές περιέχονται στο ISO “GUM” το οποίο αποτελεί τον πληρέστερο οδηγό εκτίμησης της μετρητικής αβεβαιοτήτας. Το κυριότερο πρόβλημα για να συμπεριληφθούν αποτελεσματικά οι συνδιακυμάνσης στην εκτίμηση της αβεβαιότητας είναι ο προσδιορισμός τους που απαιτεί σημαντική πειραματική εμπειρία.

Βιβλιογραφία

H. Barreto, F.M. Howland, http://www3.wabash.edu/econometrics/, 2005.
Barlow R.J., “Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences”, Wiley 1989.
Davis R.S., “Equation for the Determination of the Density of Moist Air”, Metrologia 29, (1992) 67.
Draper N.R., Smith H., “Applied Regression Analysis”, 3rd. ed., Wiley Interscience 1998.
Gottfried B.S., “Spreadsheet Tools for Engineers: Excel 97 version”, McGraw-Hill International Editions 1998.
International Organization for Standardization (ISO), “Guide to the Expression of Uncertainty in Measurements”, Geneva, 1st edition (1995).
JCGM, “Evaluation of Measurement Data - Supplement 1 to the “GUM” – Propagation of Distributions Using a Monte Carlo Method”, 1st edition JCGM 101 (2008).
Kessel R., Kacker R., 2009: “Correlation in Uncertainty of Measurement-A Discussion of State of the Art Techniques”, Proc. XIX IMEKO World Congress, Sept. 6-11 2009, Lisbon.
Kragten J., “Calculating Standard Deviations and Confidence Intervals with a Universally Applicable Spreadsheet Technique”, Analyst 119, (1994) 2161.
Park S.Y. et. al., “A Redetermination of the Ar Content of Air for Buoyancy Corrections in Mass Standard Comparisons”, Metrologia 41, (2004) 387.
Mandel J., “The Statistical Analysis of Experimental Data”, Dover Publications 1984.
Schwartz R., Borys M. and Scholz F., “Guide to Mass Determination with High Accuracy”, PTB-MA-80e 2007.
J.R. Taylor, “An Introduction to Error Analysis”, 2nd ed., University Science Books 1997.