On Choosing the Right Coordinate Transformation Method

Yaron A. FELUS, USA and Moshe FELUS, Israel

Key words: Engineering survey; GPS; Positioning, Coordinate Transformation

SUMMARY

Coordinate transformation is an important procedure aims at converting data from one reference system to another using a set of control points measured in both systems. Various methods can be used to perform this vital procedure, for example: translation, conformal, affine, projective, and polynomial transformations.

These methods are different mathematically in the number of parameters that are being used in the transformation equations and in the influence that they have on the data. A number of quality indicators have been proposed to choose the right method including: the Root Mean Squared Error (RMSE) measure, invertability, uniqueness, parsimony, and conformality.

Naturally, more complex transformations with large number of parameters (sometimes with high order terms that are not linear) will yield a smaller RMSE but will introduce more distortions and deformations into the data.

In order to choose the proper transformation method, an Akaike Information Criterion (AIC) type principle is proposed. AIC is a measure of the goodness of fit of a statistical model to a certain dataset. This criterion is employing a comprehensive point of view that takes into account both the accuracy (e.g., RMSE) as well as the complexity of the transformation function (as represented by the number of parameters). The new approach is tested in a real case study to demonstrate the implementation of it in choosing the proper transformation method between the two coordinates systems in the state of Israel.

TS 4C – Geodetic Datum II   1/10
Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

FIG Working Week 2009
Surveyors Key Role in Accelerated Development
Eilat, Israel, 3-8 May 2009

On Choosing the Right Coordinate Transformation Method

Yaron A. FELUS and Moshe FELUS, USA, Israel

1. INTRODUCTION

The past decade has seen a renewed interest among surveyors, GIS experts, and remote sensing practitioners in coordinate transformation methods. This is mainly due to the evolution of the Global Positioning System (GPS) and the need to consolidate legacy data measured in an old coordinate system with high-precision GPS-based data (Chen and Hill, 2005, Schaffrin and Felus, 2008). Coordinate Transformation techniques are essential for the creation of a Coordinate Based Cadastre (CBC) where parcel maps and boundary line information have to be accurately transformed to a new unified GPS-based coordinate system (Felus, 2007). In digital photogrammetry and remote-sensing, transformations are used to convert the pixel-based image coordinates into the map-based coordinates (Mikhail et al., 2001, p. 399). Other examples for the application of coordinate transformations include Sprinsky (1987) who utilized an affine transformation to convert digitizer coordinates to map or object coordinates, Tsenkov and Gospodinov (2002) who determined 2D-tectonic movements using an affine transformation, and Hu (2003) who converted geodetic data from state plane coordinates in the North American Datum (NAD) of 1927 to state plane coordinates in NAD83.

Coordinate transformation is the mathematical procedure to establish a geometrical relationship between a source coordinate system (local or image coordinate system) and a target coordinate system (world or object coordinate system). This procedure estimates the transformation parameters using a set of control points measured in the two coordinate systems. A plethora of transformation models have been developed; these include: surface interpolation based, polynomial, projective, affine, conformal, and translation type (e.g., Wolf and Ghilani, 1997, pp. 335-356 and Gonzalez-Matesanz et al., 2006). Each method is different in the number of parameters that are being used for the transformation and, consequently, will have different effects on the coordinates being transformed.

The question of which transformation method is more suitable for a specific project is central for the applications mentioned above and has been investigated in the past by a few authors. Chen and Hill, 2005 evaluated the performance of three transformation methods for a case study in Ireland. The Three transformation methods were the Helmert conformal transformation (seven parameters), the polynomial transformation (with 35 parameters); and an interpolated grid of Δx and Δy (10 kilometers cell resolution created by a 2nd order polynomial rubber-sheeting interpolation approach). The evaluation criteria that Chen and Hill, 2005 have used included invertability, precision - as measured by the Root Mean Squared (RMS) error of the residuals, the maximum residual and the 95% of the available residuals-, uniqueness, conformality, and extensibility. The analysis of Chen and Hill, 2005 considered many aspects of the transformation and concluded that a polynomial model is the

Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

best choice; nonetheless, a combined factor that incorporates all these criteria was not suggested in that research. Gonzalez-Matesanz et al., 2006 investigated a range of coordinate transformation models to convert control points from the ED50 datum to the ETRS89 datum. There transformation procedure had two steps: The first step used a simple coordinate transformation method (5 parameters, 7 parameters, and a polynomial) and the second step utilize an interpolated surface (with Least Squares Collocations, Rubber-Sheeting, Minimum Curvature methods) to adjust the data to local distortions. Gonzalez-Matesanz et al., 2006 used various statistical measures on the residuals and selected the 7 parameter transformation followed by the Minimum Curvature interpolation to perform the task with low RMS and extensibility. Soycan 2005 tested Two-Dimensional and Three-Dimensional similarity and polynomial transformations (12 parameters in 2D and 18 in 3D) and selected the 2D polynomial because of its low RMS error.

It seems from the literature review above that the RMS error is a popular indicator for the suitability of the transformation method. Nevertheless, the RMS error may not be an ideal criterion because a transformation model with a large number of parameters will normally yield a smaller RMS error. For example, with six control points, the 2nd degree polynomial transformation is unique and will result, theoretically, in a perfect fit (RMSE=0). Nonetheless, a transformation model with a large number of parameters is highly sensitive to outliers and may incorrectly distort, stretch, or alter the system (Kampmann, 1996). This paper presents the Akaike Information Criterion (AIC) type principle as a measure of the goodness of fit of the transformation model to a certain dataset.

The rest of the paper is organized as follows: The next section provide basic review about coordinate transformation techniques. This is followed by a presentation of AIC criterion and new approach to selecting the optimal transformation model. Then, a case study which employed the new approach to select the best transformation method for CBC in the south of Israel is described. A discussion about the limitations of the presented techniques concludes this report.

2. A REVIEW OF COORDINATE TRANSFORMATION PRODCURE

For clarity of discussion, the formulas of the similarity transformation (four-parameter transformation) will be reviewed. The transformation equation is given by:

$$\begin{bmatrix} X_T \\ Y_T \end{bmatrix} = s \cdot \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \cdot \begin{bmatrix} X_o \\ Y_o \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \cdot \begin{bmatrix} X_o \\ Y_o \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} \tag{1}$$

This transformation employs four physical parameters: s is the scale, β is the rotation, c and d are the translations along the X and Y-axes, respectively. X_T, and Y_T are the transformed coordinates (target coordinate system) while X_o, and Y_o are the original coordinates in the old coordinate system. The physical parameters $s \cdot cos(\beta)$ and $s \cdot sin(\beta)$ are often replaced by the mathematical parameters a and b as presented in the right side of Equation (1). The

TS 4C – Geodetic Datum II   3/10
Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

FIG Working Week 2009
Surveyors Key Role in Accelerated Development
Eilat, Israel, 3-8 May 2009

transformation parameters are computed from a redundant set of *n* points by using a least-squares adjustment procedure and the following system:

$$
\begin{bmatrix} x_{T1} \\ y_{T1} \\ \vdots \\ x_{Tn} \\ y_{Tn} \end{bmatrix} \approx \begin{bmatrix} x_{o1} & y_{o1} & 1 & 0 \\ y_{o1} & -x_{o1} & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{on} & y_{on} & 1 & 0 \\ y_{on} & -x_{on} & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \tag{2}
$$

$$\underbrace{}_{y} \qquad \underbrace{}_{A} \qquad \underbrace{}_{\xi}$$

This system forms the classical Gauss-Markov model:

$$
\mathbf{y} - \mathbf{e} = \mathbf{A} \cdot \xi; \qquad rank(\mathbf{A}) = 4 < 2n; \quad \mathbf{e} \sim N(0, \sigma_0^2 \cdot \mathbf{P}^{-1}) \tag{3}
$$

where **y** is a *2n×1* observation vector; **A** is a *2n×4* data matrix; ξ is a *4×1* vector of unknown parameters; **e** is a *2n×1* error or noise vector assumed to be normally distributed with zero mean; σ_0^2 is the a-priori variance component; and **P** is the *2n×2n* weight matrix.

A standard least-squares adjustment method can be employed to estimate the vector ξ of the Gauss-Markov model in (3) and compute the vector of residuals **e** as follows:

$$
\hat{\xi} = (\mathbf{A}^{\mathbf{T}} \mathbf{PA})^{-1} \mathbf{A}^{\mathbf{T}} \mathbf{Py} \tag{4}
$$

$$
\tilde{\mathbf{e}} = \mathbf{y} - \mathbf{A}\hat{\xi}
$$

where the ^ symbol denotes an estimate of a fixed variable and the ~ symbol denotes a prediction of a random variable. For simplicity, it is assumed that the observations are independent of each other and, therefore, the weight matrix **P** is a diagonal matrix, namely **P** $=Diag(p_1, p_2,...,p_{2n})$. Thus, the normalized weight matrix is defined by:

$$
\overline{\mathbf{P}} := (1/trace(\mathbf{P})) \cdot \mathbf{P} \tag{5}
$$

where the *trace* operation sums the diagonal elements of the matrix. Finally, the Weighted Sum of Squared Residuals (WSSR) is given by:

$$
WSSR := \tilde{\mathbf{e}}^{\mathbf{T}} \cdot \overline{\mathbf{P}} \cdot \tilde{\mathbf{e}}^{\mathbf{T}} \tag{6}
$$

The similarity transformation has four parameters as presented by Equations (1) and (2). Transformations with a different number of parameters, such as the translation (with two translation parameters), affine (with six parameters), projective (with eight parameters), or polynomial can be used, as well. Thus, a central question in the analysis of the coordinate

Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

FIG Working Week 2009
Surveyors Key Role in Accelerated Development
Eilat, Israel, 3-8 May 2009

transformations process is which model to use for making inferences from the data: translation, similarity, affine, projective, or polynomial transformations?

3. AKAIKE'S INFORMATION CRITERION

The problem of selecting the "best" transformation model can be solved using the Akaike's Information Criterion (AIC) (Akaike, 1974). The AIC replaces previous methods that rely on hypotheses testing to select the model that statistically minimizes the "distance" between the "true" (i.e., observed data) and the theoretical model. This distance is known as the *Kullback–Leibler divergence*. If normally distributed errors are assumed, then the AIC criterion is given by:

$$AIC := n \cdot \log(WSSR) + 2 \cdot k \tag{7}$$

where *n* is the number of observations, W*SSR* is the weighted sum of squared residuals calculated by Equation (6), and *k* is the number of parameters. The first term on the right side of equation (7) is a measure of lack of model fit, while the second term (i.e., *2·k*) can be interpreted as a "penalty" for increasing the size of the model (this penalty enforces parsimony in the number of parameters). Hence, a good mathematical model is one that has the smallest AIC score.

The AIC may perform poorly in small datasets (Burnham and Anderson, 2002). Consequently, a refined Akaike information criterion, which we shall denote by AICc, was developed for model selection in small datasets, as follows:

$$AICc := n \cdot \log(WSSR) + 2 \cdot k \cdot (n/(n-k-1)) \tag{8}$$

The refined criterion – AICc – should be used when the ratio *n/k* is small (e.g., < 40).

4. CASE STUDY

An important step in the cadastral data enhancement workflow is a global transformation from the old Israeli coordinate system (Cassini-Soldner projection) to the new Israeli coordinate system (Transverse Mercator projection). Due to large local distortions, this transformation process has to be performed empirically, using national control points with values in both systems.

This elegant statistical tool was employed to select the optimal transformation model in converting a set of parcel maps (i.e., cadastral blocks 1928, 1957, 1954, 1953) from the old Israeli coordinate system to the new one. Four transformation models were tested using 19 control points in the project area. The results of the experiments are summarized in Table 1. These results show that a similarity transformation is the most suitable model for this specific area, with the smallest AICc value of -51.891. Note how the weighted sum of squared residuals (i.e., *WSSR* value in Table 1) decreases as the number of parameters in the transformation is increased.

Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

Table 1. *Comparing different transformation models*

	Translation	Similarity	Affine	Projective
Number of Parameters (k)	2	4	6	8
WSSR	0.9524	0.544	0.491	0.4519
AICc	-48.103	**-51.891**	-45.176	-34.323

The refined Akaike information criterion (AICc) may also provide some guidance in eliminating substandard information and selecting the optimal set of control points for a transformation in a given area. Different methods and techniques have been proposed to identify and remove erroneous data and outliers. These methods are often subjective and depend on the statistical rejection criteria (cf. Koch,1999 or Baarda, 1968). In this study, we employed the AICc as an objective criterion to evaluate the consistency of a dataset. The area of Israeli parcel map 1928 which includes nine control points was selected, and a similarity transformation was performed in eight experiments as presented in Table 2. In every experiment, the point with the largest *residual-distance* was removed from the list. The residuals-distance was calculated using the formula: $d_{e_i} = \sqrt{(e_{x_i}^2 + e_{y_i}^2)}$, where e_{x_i} and e_{y_i} are the residuals in the *X* and *Y* directions of control point *i* as extracted from vector $\tilde{e}$ in equation (4). Moreover, in every experiment the following quantities were computed: the transformation parameters, residuals and the residual-distance, AICc value, weighted sum of squared residuals (*WSSR*), and the diagonal elements of the Hat matrix which provide an indication on the relative influence of every point on the transformation (Kampmann, 1996). The Hat matrix – H – is defined by:

$$\mathbf{H} := \mathbf{A} \cdot (\mathbf{A}^T \cdot \mathbf{P} \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{P} \tag{9}$$

A large diagonal value on the Hat matrix suggests a point with a large influence (leverage point), while a small diagonal value suggests an insignificant point.

Table 2. *Selecting a consistent set of control points using the AICc. Note that the smallest AICc (-74.684) was obtained after removing point 40-J from the list*

Experiment	1	2	3	4	5	6	7	8
No. of points	9	8	7	6	5	4	3	2
WSSR	0.5445	0.0726	0.0535	0.0404	0.0250	0.0137	0.0073	0
AICc	-51.891	**-74.684**	-65.487	-54.599	-43.893	-29.605	7.7833	∞
Point removed	-	**40-J**	44-J	41-J	34-K	59-J	152-G	106-NG

Figure 1 presents the spatial configuration of the nine control points, and Table 4 presents the residuals, the residual-distance, and the diagonal elements of the Hat matrix in the first experiment. Note that new residuals had to be calculated at every experiment.

TS 4C – Geodetic Datum II   6/10
Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

FIG Working Week 2009
Surveyors Key Role in Accelerated Development
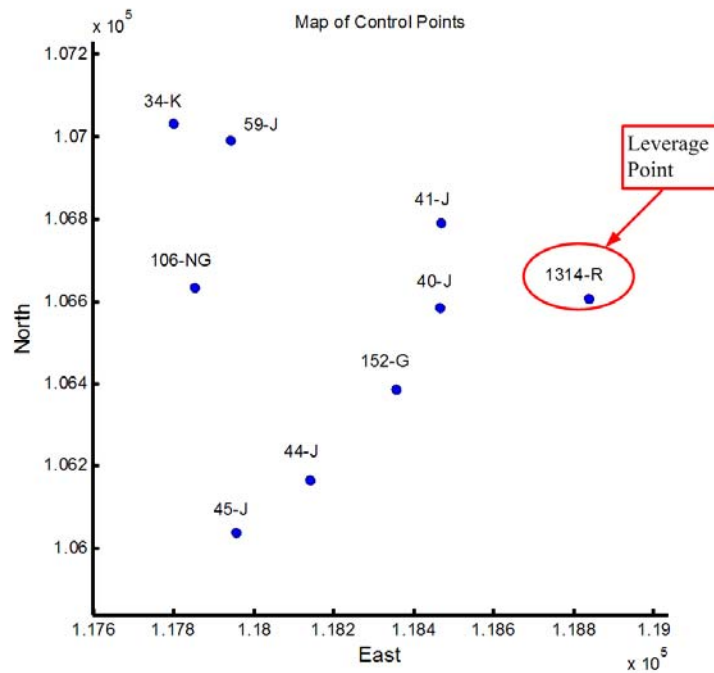Eilat, Israel, 3-8 May 2009

Fig. 1. A diagram of control points. Point 1314-R was identified as a leverage point because it has the largest diagonal value on the Hat matrix; point 152-G is the least significant point since it has the lowest diagonal value on the Hat matrix.

The AICc values in Table 2, suggest that point 40-J should be omitted to create a consistent dataset with a low AICc value of -74.68. Note that this point has a low influence on the transformation, as being indicated by the Hat matrix (but not the lowest). Moreover, this point (40-j) was not identified as an outlier using standard outlier detection tests such as (Baarda, 1968) (with $\alpha = 0.05$) or the Chebyshev criterion which excludes points with residuals outside the range of (*Mean* $\pm$ 3 $\times$ standard deviation). The last two rows in Table 3 provide information about the mean and the standard deviation of the residuals.

TS 4C – Geodetic Datum II
7/10
Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

FIG Working Week 2009
Surveyors Key Role in Accelerated Development
Eilat, Israel, 3-8 May 2009

Table 3. *The residuals, residual-distance, and the Hat matrix values in the first experiment. Note that point 40-j has the largest residual-distance, and point 1314-R has the largest diagonal value of the Hat matrix (i.e., a leverage point)*

Point Name	Residual X (e_x)	Residual Y (e_y)	Residual-distance $d_{e_i} = \sqrt{(e_{x_i}^2 + e_{y_i}^2)}$	Diagonal of the Hat matrix
152-G	-0.154	-0.075	0.171	**0.0326**
44-J	0.025	0.030	0.039	0.0924
45-J	-0.138	0.019	0.140	0.1856
59-J	0.027	-0.053	0.059	0.1236
106-NG	-0.128	-0.023	0.130	0.0657
34-K	0.035	-0.123	0.128	0.1911
40-J	0.423	0.472	**0.634**	0.0362
41-J	-0.112	-0.199	0.229	0.0601
1314-R	0.022	-0.049	0.053	**0.2122**
Mean	0.0001	-0.00011	0.1758	-
St. Dev.	0.1780	0.1904	0.1823	-

5. CONCLUSIONS AND FURTHER RESEARCH

The refined Akaika Information Criterion (AICc) was used to select the appropriate coordinate transformation model. Other criterion, such as the Mallows statistic Cp (Mallows, 1973), may be used as well. The Mallows statistic is given by:

$$Cp = (\hat{\sigma}^2)^{-1} \cdot (WRSS) - n + 2 \cdot k \qquad (10)$$

where $\hat{\sigma}^2$ is a properly chosen estimate of the posteriori reference variance, n is the number of observations, and k is the number of parameters. However, the Mallows statistic Cp was criticized as being subjective to the choice of the posteriori reference variance (Akaike, 1974) and, therefore, was not used in this research.

Another important criteria is the Bayes Information Criterion (Schwarz, 1978).

$$BIC := n \cdot \log(WSSR) + \log(n) \cdot k \qquad (11)$$

Similar to the AIC, the model with the lower value of BIC is the one to be preferred. The BIC is an increasing function of RSS and an increasing function of k. Finally the Kashyap's criterion (Kashyap, 1982) is a modification of the BIC with additional element added to equation (11). Further research is needed to compare the different criteria AIC, BIC and Kashyap's and identify the one most suitable in mapping applications.

Applications of the AICc to eliminate erroneous data are still under investigation although the initial study which was presented in Table 2 demonstrated how it can be performed. The

presented coordinate transformation applications can be improved by using the Total Least-Squares (TLS) approach with the Error-In-Variables (EIV) model as described in Schaffrin, and Felus (2008). The use of a EIV model vs Gauss-Markov model depend on the problem at hand and the measured variables.

REFERENCES

1. Akaike, H. (1974) A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, Vol. 19, No. 6: 716-723.
2. Baarda, W., 1968. A Testing Procedure for Use in Geodetic Networks. Publications on Geodesy 9 (Vol. 2, No. 5), Delft, Netherlands, 97 p.
3. Burnham, K.P. and Anderson, D.R. (2002) Model Selection and Multimodel Inference: A Practical-Theoretic Approach, 2nd Ed. Springer, New York, 496 p.
4. Chen W, Hill C (2005) Evaluation procedure for coordinate transformation. Journal of Surveying Engineering 131(2): 43-49
5. Felus Y .A. (2007) On the positional enhancement of digital cadastral maps. Survey Review, Vol 39 (306), pp. 268-281.
6. Gonzalez-Matesanz J., Dalda A., Malpica J.A.(2006) A range of ED50-ETRS89 datum transformation models tested on the Spanish geodetic network. Survey Review, Vol 38 (302), pp. 654-667.
7. Kampmann, G., 1996 New Adjustment Techniques for the Determination of Transformation Parameters for Cadastral and Engineering Purposes. *Geomatica*, Vol. 50, No. 1: 27-34.
8. Kashyap, R. L. (1982). Optimal Choice of AR and MA Parts in Autoregressive Moving Average Models, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 4, 99-104.
9. Koch, K.-R., 1999. Parameter Estimation and Hypothesis Testing in Linear Models (2nd edition). Springer-Verlag, Berlin, 333 p.
10. Mallows, C.L., 1973. Some comments on C.P. Technometrics, Vol. 15, No. (4): 661-675.
11. Mikhail EM, Bethel JS, McGlone JC (2001) Introduction to Modern Photogrammetry. Wiley, New York/Chichester, 479 pages
12. Schaffrin, B. and Felus, Y. A. (2008) On the Multivariate Total Least-Squares Approach to Empirical Coordinate Transformations: Three Algorithms. Journal of Geodesy Vol 82(6), pp. 373-383.
13. Schwarz, G. (1978). Estimating the Dimension of a Model, Annals of Statistics, Vol. 6, 461-464.
14. Soycan, M., (2005) Polynomial versus Similarity Transformations between GPS and Turkish Reference Systems. Survey Review, Vol. 38, No. 295: 58-69.

TS 4C – Geodetic Datum II   9/10
Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

FIG Working Week 2009
Surveyors Key Role in Accelerated Development
Eilat, Israel, 3-8 May 2009

15. Sprinsky WH (1987) Transformation of positional geographic data from paper-based map products. The American Cartographer 14(4): 359-366

16. Tsenkov T, Gospodinov S (2002) Determination of 2D-tectonic deformations using affine transformation. Zeitschrift für Vermessungswesen 127(4): 262-264

17. Wolf PR, Ghilani CD (1997) Adjustment Computations: Statistics and Least-Squares in Surveying and GIS. Wiley, New York, 564 pages

BIOGRAPHICAL NOTES

Dr. **Yaron A. Felus**
holds a BS from Ben-Gurion University, Israel (Electrical and Computer Engineering), a M.S. from the International Institute for Geo-Information Science and Earth Observation, Netherlands (Geomatics), and both M.S. and Ph.D. from the Ohio State University (geodetic science). He is a Professional Surveyor in Michigan, USA and a Licensed Surveyor in Israel.

Mr. **Moshe Felus**
a Licensed Surveyor in the state of Israel, is the founder and manager of Halperin-Felus Surveying and Photogrammetry Ltd. He has more than 35 years of experience in surveying engineering and cadastral work for planning and construction.

CONTACTS

Yaron A. Felus
Associate Professor
Surveying Engineering Department
Ferris State University
USA
E-mails: felusy@ferris.edu

Moshe Felus
Halperin-Felus Surveying and Photogrammetry Ltd.
10 Hachsarat Hayeshuv St,
Rishon LeZion, 75612,
ISRAEL
Tel: + 972 3 9627082
E-mail: moshe@hf-mapping.co.il
Web: http://www.hf-mapping.co.il

TS 4C – Geodetic Datum II   10/10
Yaron A. Felus and Moshe Felus
On Choosing the Right Coordinate Transformation Method

FIG Working Week 2009
Surveyors Key Role in Accelerated Development
Eilat, Israel, 3-8 May 2009