



Initial results of sediment yield measurement interpretation using a regional approach: Southern Italy case study

Paolo Porto^{1,2} and Giovanni Callegari²

¹Department of Agraria, University Mediterranea of Reggio Calabria, Reggio Calabria, 89122, Italy

²CNR – Istituto per I Sistemi Agrari e Forestali per il Mediterraneo,
Sezione Ecologia e Idrologia Forestale, Rende (Cs), Italy

Correspondence: Paolo Porto (paolo.porto@unirc.it)

Published: 1 August 2019

Abstract. The theoretical TCEV (Two Component Extreme Value) distribution was applied to interpret the sediment yield datasets available in Southern Italy. The analysis is based on hydrological data collected for twelve catchments located in Calabria and Basilicata. A hierarchical approach was used to obtain a regional parent distribution which was used to determine the return time for each event. The hierarchical approach proposed in this study includes two stages. The first stage served for calibration and made it possible to estimate the parameters of the theoretical TCEV distribution. More specifically, the hypothesis of homogeneity with regard to the skewness coefficient and the coefficient of variation was verified using the datasets related to nine catchments. The second stage consisted in verifying the goodness of the theoretical distribution on three independent datasets provided by three experimental catchments not involved in the calibration. Overall results show that, even if the TCEV distribution was conceived to estimate peak flow, its concept of “double component” can be extended to predict sediment yield on a regional scale.

1 Introduction

Prediction of soil erosion and sediment yield, corresponding to a fixed return time, is necessary to adopt soil conservation strategies for reducing costs of on-site and off-site impacts in areas at risk.

In Southern Italy, high rates of soil erosion and sediment yield are well documented both on cultivated land (Porto and Walling, 2012) and in areas covered by forests (Porto et al., 2004). Over the last 60 years, predictions of sediment yield are based both on the use of traditional monitoring techniques, such as observations on experimental plots or catchments, and on different types of models that vary from empirical-parametric approaches, such as SEDD (Ferro and Porto, 2000) or revised versions of the USLE (Cinnirella et al., 1998) to more recent physically-based models, such as WEPP (Nearing et al., 1989), whose goal is to simulate both the detachment and transport of soil particles.

However, even if these models provide good estimates of long-term average values of sediment yield, they cannot make reliable predictions associated to a fixed return time especially if they are associated with extreme events.

A recent study (Porto and Callegari, 2019), carried out on three experimental catchments (W1, W2, and W3) located in Calabria (Southern Italy), demonstrated that a significant proportion of the annual sediment yield in these catchments is due to the extreme event occurred each year which produced more than 50 % of the total annual sediment yield with a very few exceptions.

Looking at the frequency distribution of these measurements of sediment yield, the same authors showed that the empirical distribution of the normalized variable $x_i = SY_{MAX,i}/\mu(SY)$, where $SY_{MAX,i}$ is the maximum annual value and $\mu(SY)$ is the mean annual value of the entire monitoring period, is skewed. In particular, results of observation on the catchment W2 (see Porto and Callegari, 2019) shows that two components can be identified (Fig. 1). One of them

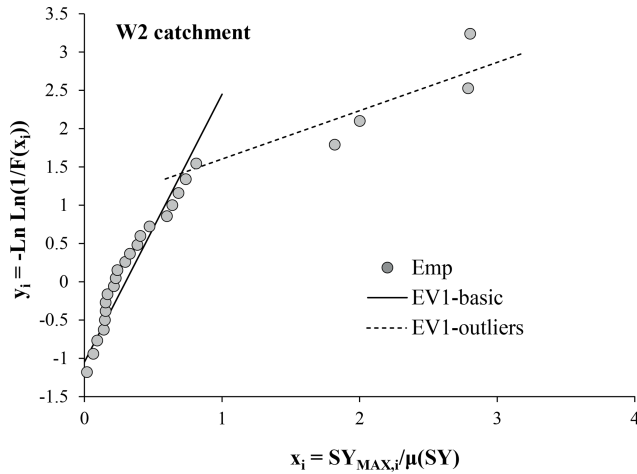


Figure 1. Gumbel's plot for the dataset of the normalized variable, $x_i = SY_{MAX,i} / \mu(SY)$.

is associated with ordinary events (basic component) and a second one is characterized by extreme events (outliers component).

The authors concluded that the process can be represented by two Extreme Value Type 1 (EV1) independent distributions, one for each component, capable of associating each event with a corresponding return time T . The double component indicated in Fig. 1 is a well known effect when studying annual records of maximum peak flow values (see Rossi et al., 1984) and a more sophisticated approach can be used to explain it.

This approach is based on the TCEV (Two Component Extreme Value) model whose ability was well demonstrated in predicting peak flow and critical rainfall corresponding to a fixed return time (see Rossi et al., 1984; Ferro and Porto, 1999, 2006). The TCEV distribution is based on the assumption that in each historical sample of the observed variable x , one or several annual maximum values are significantly higher than the bulk of the remaining data, so that two components are clearly distinguishable.

Therefore, the theoretical CDF of the TCEV law $F(x)$ can be expressed as a probabilistic model with four parameters derived from the product of two EV1 distributions:

$$F(x) = \exp \left[-\lambda_1 \exp \left(-\frac{x}{\theta_1} \right) - \lambda_2 \exp \left(-\frac{x}{\theta_2} \right) \right] \quad (1)$$

where: λ_1, λ_2 are the shape parameters and represent the mean number of events which belong respectively, to the basic (component 1) and the outlying (component 2) components; and θ_1, θ_2 are the scale parameters, which represent the at-site central values of the x variable for each component.

The basic component is characterized by a high number of events and by values of x less than those corresponding to the outlying component so that $\lambda_1 \gg \lambda_2$ and $\theta_1 \ll \theta_2$. The

main advantage of the TCEV model, if compared to the EV1 distribution, is based on the fact that the former can be used at a regional scale because their parameters (or their combination) showed to be scale invariant for large areas.

In order to apply the model at a regional scale it is useful to introduce the standardized variable $y = \left(\frac{x}{\theta_1} \right) - \lambda_1$ by which Eq. (1) becomes:

$$F(y) = \exp \left[-\exp(-y) - \Lambda_* \exp \left(-\frac{y}{\Theta_*} \right) \right] \quad (2)$$

where: $\Theta_* = \theta_2 / \theta_1$ and $\Lambda_* = \lambda_2 / \lambda_1^{1/\Theta_*}$.

Introducing the dimensionless variable x' , equal to the ratio between x and the mean value μ of the TCEV distribution, Eq. (2) becomes:

$$F(x') = \exp \left\{ -\lambda_1 [\exp(\alpha)]^{-x'} - \Lambda_* \lambda_1^{1/\Theta_*} \left[\exp \left(\frac{\alpha}{\Theta_*} \right) \right]^{-x'} \right\} \quad (3)$$

named growth curve, in which:

$$\alpha = \frac{\mu}{\theta_1} = 0.5772 + \ln \lambda_1 - \sum_{j=1}^{\infty} \frac{(-1)^j \Lambda_*^j \Gamma \left(\frac{j}{\Theta_*} \right)}{j!} \quad (4)$$

where $\Gamma()$ is the gamma function.

For fixed return period T , Eq. (3) requires only 3 parameters (Λ_* , Θ_* and λ_1) for the calculation of the variable x' into a selected area (i.e. conversion of growth curve to quantile). It is assumed that the first two parameters, Λ_* and Θ_* , are constant over a very large area and the parameter λ_1 is considered constant for each sub-region into which the region is divided. When the three parameters of Eq. (3) are defined and taking into account of the definition of the dimensionless variable x' , the corresponding value of the variable x can be obtained by the following equation:

$$x = x' \mu \quad (5)$$

where: the variable μ (or scale parameter) can be replaced by the empirical mean of the single records.

Considering the logical dependence of sediment yield on the magnitude of peak flow (see among the others Williams and Berndt, 1977), the aim of this work is to interpret the observed values of sediment yield using an approach based on the TCEV distribution.

2 Materials and methods

The Italian Hydrographic Service (SIMI) has published long-term measurements of daily and monthly suspended sediment load (SSL) on many catchments throughout the country. In the present contribution, monthly measurements of SSL, which were collected on 11 rivers located in Calabria, and Basilicata, have been used for analysis. This step required also a preliminary aggregation of the data to generate

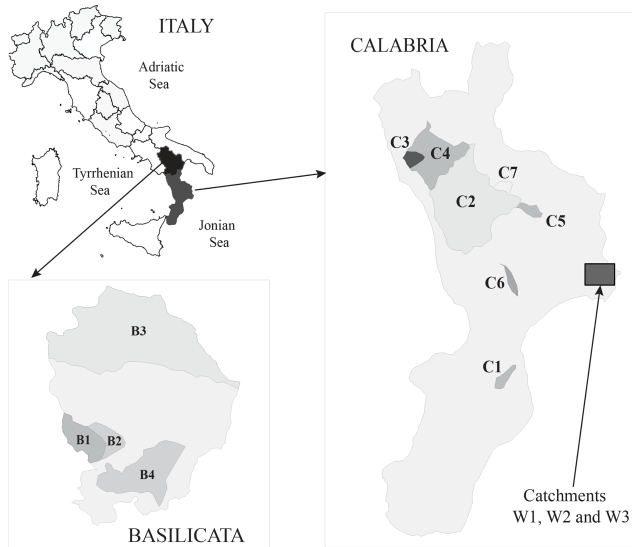


Figure 2. Location of the catchments investigated in this study.

the annual datasets of sediment yield (see Table 1 and Fig. 2 for details).

These datasets have been integrated by 3 long-term records of sediment yield obtained in three experimental catchments (W1, W2, and W3) located in Calabria (see Porto et al., 2004). In this case, the original measurements, collected at event scale, have been aggregated at annual scale to make them comparable with the previous datasets.

More specifically, the 11 datasets were used to derive the growth curve (Eq. 3) for the entire region, here assumed as a unique homogeneous area, while the three catchments W1, W2, and W3 were used to test the validity of the Eq. (3) on independent datasets located in the same region.

In other words, the hierarchical approach proposed in this study consists of two levels: (1) the first level aimed at calculating the 3 parameters (Λ_* , Θ_* and λ_1) of the Eq. (3) on the area where the catchments are located. In this respect, a comparison between the empirical and theoretical CDFs of γ was performed; (2) the second level served to validate the Eq. (3) on the three independent datasets for which the empirical CDFs of the variable x are available.

3 Results and discussions

3.1 Model calibration

The analysis carried out at the first level required the basic assumption that the entire region can be considered homogeneous with respect to the skewness coefficient, γ and the coefficient of variation, CV (see Fiorentino et al., 1987; Beran et al., 1986). However, considering the short length of some historical datasets ($N < 10$ years), some preliminary adjustments of these records were necessary. More specifically, based on geomorphic similitude, the two datasets of

Table 1. Details of stream gauges and rivers considered in this study.

River	Station	Code	Years	N	Area (km ²)
BASILICATA					
AGRI	Grumento	B1	1962–1972	10	278.0
AGRI	Tarangelo	B2	1968–1972	5	507.0
BRADANO	Tavole Palatine	B3	1933–1971	34	2743.0
SINNI	Valsinni	B4	1938–1973	26	1142.0
CALABRIA					
ALACO	Mammone	C1	1961–1979	18	14.8
CRATI	Conca	C2	1933–1970	31	1332.0
ESARO	Cameli	C3	1964–1971	8	55.4
ESARO	La Musica	C4	1960–1969	9	532.0
LESE	Schiena d'Asino	C5	1953–1980	19	60.0
MELITO	Olivella	C6	1962–1977	14	41.2
TRIONTO	Difesa	C7	1964–1979	15	31.7

River Agri (see Table 1) were bulked into one single dataset with $N = 15$; the same assumption was made for the two records of River Esaro for which a final single dataset with $N = 17$ was obtained.

As explained above, in the first step of this analysis, the TCEV model required the estimation of the regional parameters Λ_* and Θ_* referred to these new nine series. This analysis was conducted using the procedure suggested by Fiorentino and Gabriele (1985) that led to the following ML estimates of the two parameters:

$$\Lambda_* = 0.675; \Theta_* = 11.574 \quad (6)$$

In order to verify the homogeneity of the entire region with regard to γ , Monte Carlo experiments were performed for generating theoretical sequences having a fixed size N and distributed according to a TCEV model with regional parameters given by Eq. (6).

The CDF of the empirical skewness related to the 9 samples and that generated by the regional model are compared in Fig. 3. The theoretical CDF was obtained using a Monte Carlo technique based on 20 000 synthetic series having a sample size N equal to the mean value calculated from the empirical datasets.

Even if a perfect overlapping cannot be expected considering the limited number of datasets, Fig. 3 shows a certain ability of the TCEV distribution to reproduce the empirical skewness CDF.

The second step of the analysis is based on the assumption that the investigated region should be divided into smaller homogeneous areas, named sub-regions, in which the λ_1 parameter can be assumed constant. Because the λ_1 parameter is related to CV by the following equation:

$$CV = \frac{0.577}{\log \Lambda_1 + 0.251} \quad (7)$$

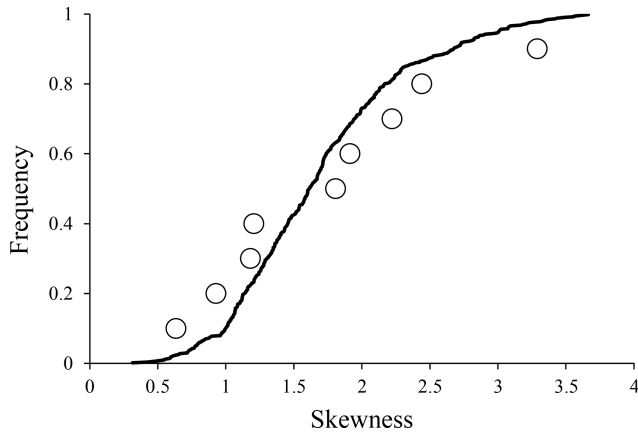


Figure 3. Comparison between the empirical CDF of the skewness coefficient calculated from the nine datasets and the regional CDF obtained from the theoretical distribution.

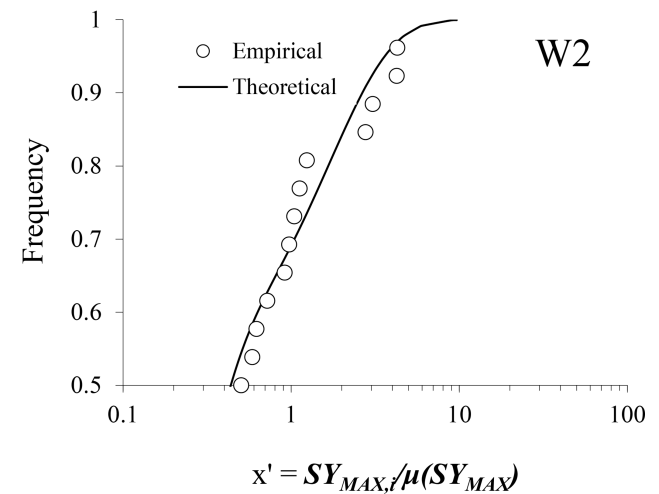
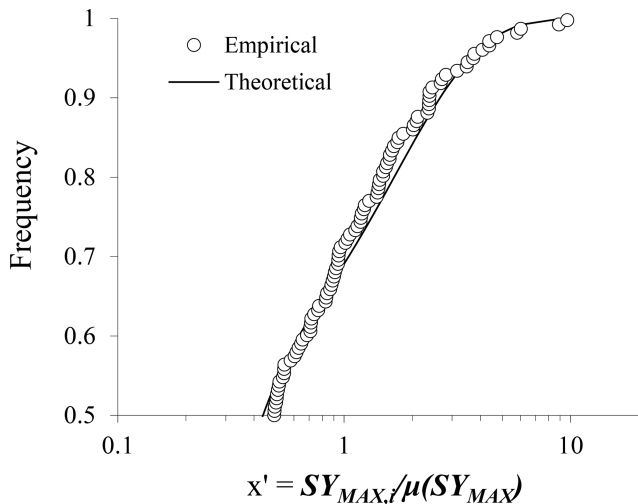
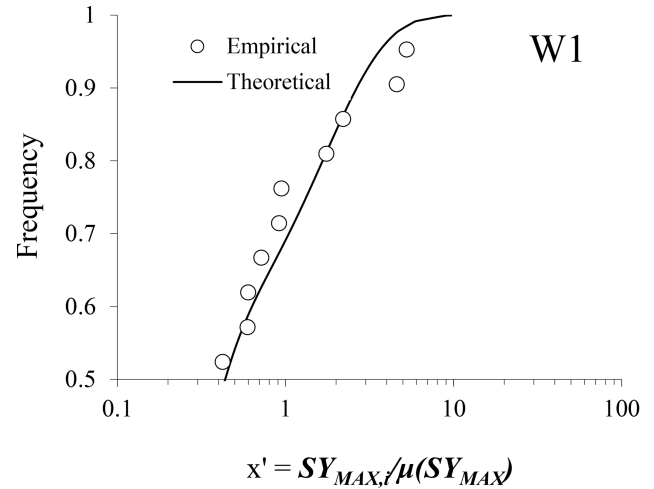


Figure 4. Comparison between empirical and theoretical CDF of the variable x' for the 9 investigated datasets.

this is equivalent to defining sub-regions within which the coefficient of variation of the individual series can be assumed constant. Unfortunately, the number of series available in the entire region (9 datasets) is not enough to proceed with a further geographic subdivision and this hypothesis could not be considered. In the light of this limitation, the analysis was carried out considering a unique homogeneous region even at the second hierarchical level. Consequently, according to Arnell and Gabriele (1988), the value of Λ_1 was determined using the TCEV distribution with the regional values of Λ_* and Θ_* given by Eq. (6). Using the data obtained from the historical datasets, the parameters Λ_1 and α assume the following values:

$$\Lambda_1 = 5.974; \alpha = 8.776 \quad (8)$$

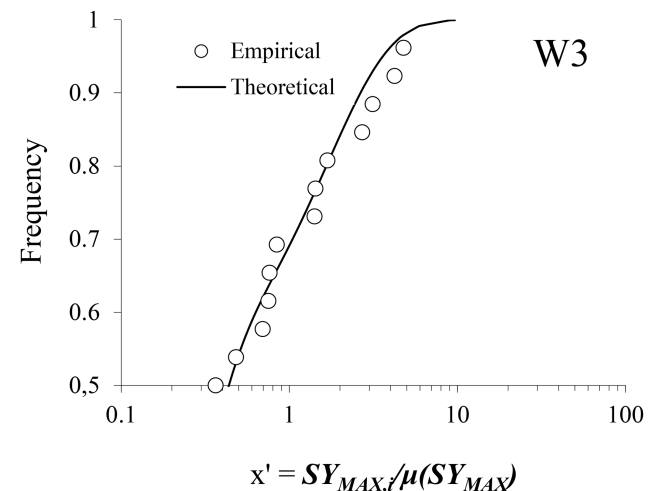


Figure 5. Comparison between empirical and theoretical CDFs of the variable x' for the 3 independent datasets (W1, W2 and W3).

Figure 4 shows that the growth curve calculated by the Eq. (3), with the parameters given by the Eqs. (6) and (8) fits very well the higher values ($f(x') > 0.5$) of the x' empirical distribution suggesting that the hypothesis of unique homogeneous region is plausible.

3.2 Model performance

The regional growth curve given by the Eq. (3) with the calibrated values of the 4 parameters (Λ_* , Θ_* , λ_1 and α) was tested on three independent datasets available for three experimental catchments (W1, W2, and W3) located in Calabria (see Fig. 2). Details of these catchments are reported elsewhere (see Ferro and Porto, 2000). Here, only the CDFs of the empirical variable x' are considered and the comparison with the theoretical distribution is reported for each catchment in Fig. 5.

The visual inspection of the three graphs in Fig. 5 suggests that, the model fits reasonably well the higher values ($f(x') > 0.5$) of the x' empirical distributions suggesting that the regional growth curve is very robust and can be used for obtaining reliable predictions even for large events.

4 Conclusions

Prediction of sediment yield generated by large events is very important in areas not covered by monitoring stations or for the areas with limited observational periods. Based on a previous contribution by Porto and Callegari (2019), the study presented herein reports a regional investigation carried out in Southern Italy aimed at evaluating the frequency distribution of sediment yield. The procedure applied in this contribution is based on the use of the TCEV model and its hierarchical approach to get reliable estimates of the regional parameters of the distribution. The choice is due to the ability of the TCEV distribution to reproduce the skewed empirical distributions of sediment yield observed in many historical datasets and to generate estimates of sediment yield associated with a fixed return time. Overall, the study demonstrated that even if the TCEV distribution was conceived to provide estimates of peak flow, its basic assumptions and its hierarchical procedure can be used to predict sediment yield at regional scale.

Data availability. Data collected by the Italian Hydrographic Service (SIMI) can be downloaded from the website <http://www.gruppoalluvioni.it/annali-idrologici/> (last access: 1 August 2018). Data related to the catchments W1, W2 and W3 are available upon request to the authors.

Author contributions. All authors setup the research, analyzed the results and participated in writing the paper.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue “Land use and climate change impacts on erosion and sediment transport”. It is a result of the ICCE Symposium 2018 – Climate Change Impacts on Sediment Dynamics: Measurement, Modelling and Management, Moscow, Russia, 27–31 August 2018.

Acknowledgements. The study has been finalized in the frame of the Erasmus + KA2 – Cooperation for innovation and the exchange of good practices – Capacity Building in the field of Higher Education – Soil Erosion and Torrential Flood Prevention: Curriculum Development at the Universities of Western Balkan Countries/SETOF.

Financial support. This research has been supported by the European Project SETOF (grant no. 598403-EPP-1-2018-1-RS-EPPKA2-CBHE-JP).

References

- Arnell, N. W. and Gabriele, S.: The performance of the two component extreme value distribution in regional flood frequency analysis, *Water Resour. Res.*, 24, 879–887, 1988.
- Beran, M., Hosking, J. R. M., and Arnell, N. W.: Comment on ‘Two component extreme value distribution for flood frequency analysis’, by Rossi et al., *Water Resour. Res.*, 22, 263–266, 1986.
- Cinnirella, S., Iovino, F., Porto, P., and Ferro, V.: Antierosive effectiveness of Eucalyptus coppices through the cover management factor estimate, *Hydrol. Process.* 12, 635–649, 1998.
- Ferro, V. and Porto, P.: Regional analysis of rainfall-depth-duration equation for South Italy, *J. Hydrol. Eng.*, 4, 326–336, 1999.
- Ferro, V. and Porto, P.: Sediment Delivery Distributed (SEDD) Model, *J. Hydrol. Eng.*, 5, 411–422, 2000.
- Ferro, V. and Porto, P.: Floods frequency analysis for Sicily (South Italy), *J. Hydrol. Eng.*, 11, 110–122, 2006.
- Fiorentino, M. and Gabriele, S.: Distribuzione TCEV: metodi di stima dei parametri e proprietà statistiche degli stimatori, *Geodata* n. 25, Cosenza, 1985 (in Italian).
- Fiorentino, M., Gabriele, S., Rossi, F., and Versace, P.: Hierarchical approach for regional flood frequency analysis. *Regional Flood Frequency Analysis*, edited by: Singh, V. P., D. Reidel Publishing Company, Springer, the Netherlands, 2, 39–49, 1987.
- Nearing, M. A., Foster, G. R., Lane, L. J., and Finkner, S. C.: A process-based soil erosion model for USDA-Water Erosion Prediction Project Technology, *Transactions of the ASAE*, 32, 1587–1593, 1989.
- Porto, P. and Callegari, G.: A Probabilistic Approach to Interpret Long-Term Observations of Sediment Yield in Experimental Catchments in Southern Italy, in: *Climate Change Impacts on Sediment Dynamics: Measurement, Modelling and Management*, edited by: Chalov, S., Golosov, V., Li, R., and Tsyplenkov, A., Springer Nature Switzerland AG, Cham, Switzerland, 110–117, 2019.

- Porto, P. and Walling, D. E.: Using plot experiments to test the validity of mass balance models employed to estimate soil redistribution rates from ^{137}Cs and $^{210}\text{Pb}_{\text{ex}}$ measurements, *Appl. Radiat. Isotopes*, 70, 2451–2459, 2012.
- Porto, P., Walling, D. E., and Callegari, G.: Validating the use of caesium-137 measurements to estimate erosion rates in three small catchments in Southern Italy, *IAHS Publ.*, 288, 75–83, 2004.
- Rossi, F., Fiorentino, M., and Versace, P.: Two-component extreme value distribution for flood frequency analysis, *Water Resour. Res.*, 20, 847–856, 1984.
- Williams, J. R. and Berndt, H. D.: Sediment Yield Prediction Based on Watershed Hydrology, *Transactions of the ASAE*, 20, 1100–1104, 1977.