Zeitschrift: Schweizerische Zeitschrift für Forstwesen = Swiss foresty journal = Journal forestier suisse

Herausgeber: Schweizerischer Forstverein

Band: 148 (1997)

Heft: 10

Artikel: A new dryness index and the non-parametric estimate of forest fire probabilities

Autor: Mandallaz, Daniel / Ye, Ronghua

DOI: https://doi.org/10.5169/seals-765491

A New Dryness Index and the Non-Parametric Estimate of Forest Fire Probabilities

By *Daniel Mandallaz* and *Ronghua Ye*

Keywords: Forest fires; dryness index; prediction; non-parametric. FDK 111: 43: UDK 519.248,2

1. Introduction

In the recent years intensive studies have been done on the prediction of forest fires in the Mediterranean countries. The main objectives of these studies were firstly the comparison of various dryness indexes and the selection of the «best» ones (*Sol* 1995), and secondly the development of new indexes and methods for the forest fire prediction and for the assessment of the goodness of fit of various methods (*Mandallaz* und *Ye* 1996).

A dryness index is generally a function of meteorological variables, for instance, the wind velocity, precipitation, temperature, air humidity or the cumulative values thereof. If the index value is «high», the day can be predicted as «highly dangerous» or «dangerous», otherwise as «not dangerous». If a threshold is used, an «alarm» can be launched if the index value exceeds the threshold.

There are many dryness indexes in use, for example the Canadian indexes (IFM, IPI, etc.), the French indexes (RN, SEUL, etc.), the Italian index (IREPI), the Spanish index (ICONA) and the Portuguese index (IP). *Bovio, Sol* und *Viegas* (1994, Sol, 1995) have done an inter-comparison study of these indexes in Italy, France und Portugal.

Most of the indexes are designed on the basis of some theoretical und empirical knowledge and their construction is unfortunately relatively complicated and not transparent. The index presented in this paper, developed in the EU Project MINERVE II for the southern Alps of Switzerland, is intuitively very simple, easy to implement in practice and has a high performance in comparison with other indexes used.

In order to find out the weakness of an index the non-parametric estimate of the fire probabilities based on the index is used; this is particularly useful to avoid large conceptual mistakes in the construction of the index. Furthermore, it is possible to give a simple criterion for an index to be essentially correct. Comparison studies have shown that indexes satisfying this criterion will have higher overall performance than those which do not.

2. Construction of the index

2.1 Data material

The data material was provided by the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), and the meteorological service in Tessin. It includes the forest fire records and the meteorological data for the southern Alps of Switzerland from 1980 to 1995. The whole region is divided into three sub-regions according to the climate and vegetation type:

Region 1: southern Tessin (forest area ca. 257 km^2)
Region 2: central Tessin (forest area ca. 673 km^2)
Region 3: northern Tessin, Vallese, Grigiono italiano
(forest area ca. 850 km^2)

There were 1433 forest fires recorded, 610 in Region 1, 567 in Region 2 und 256 in Region 3. The data is analysed separately for two forest fire seasons: winter (November – May) und summer (June – October). Most forest fires occurred in winter.

For the calculation of the indexes the following meteorological data are used:

- air humidity at 12h and the daily average value
- maximum wind velocity during one hour and the daily average value
- air temperature at 12h and the daily average value
- precipitation during the last 24 hours
- soil temperature at 12h
- global sun radiation and relative sun radiation

2.2 Construction of the index

The idea for the construction of the index is very simple and intuitive: imagine a swimming pool with water. At the beginning the pool is full and it is therefore «not dry». If it does not rain, the water evaporates, the water level in the pool goes down and it becomes «dryer», and vice versa. The pool depth minus the water level in the pool is then essentially our dryness index.

We define three versions of the index:

1. ETC(day n+1) = ETC(day n) + (Ep - precipitation)
2. ET(day n+1) = ETC, but ET = 0 if Ep < precipitation
3. ETP(day n+1) = ETP(day n) + (Ep - precipitation), but
   In case of Ep < precipitation:
      on the first day: ETP = TSol*ETC(day n+1)/C, where C is a constant
      on the following days: ETP(day n+1) = ETP(day n) + TSol*ETC(day n+1)/C,
   until ETC = ETP. Only the most recent precipitation is considered (see also the appendix).

where Ep is the evaporation and is calculated according to the Penman's formula, whose accuracy is established by many studies (*Item,* 1974, 1981). The starting value of the indexes is 0 (this is the case after a heavy rain), after some days (ca. 2 months) the index values are independent of the starting value. A SAS program is given in the appendix.

The main difference between the three versions is due to the treatment of the rainy days. Version 1 treats rainy days by simply subtracting the precipitation surplus, version 2 sets the index value simply to 0 (the pool is now full), whereas the water level of version 3 lies between the former two.

For version 3 we consider the following factors which may influence the speed of the water infiltration into the soil: 1) the dryness (we use here the index ETC as indicator for dryness), the dryer it is, the faster the infiltration; 2) the soil temperature (TSol), the higher the temperature, the faster the infiltration, according to the simple formula: TSol*ETC/C.

We now determine the constant C by considering the boundary conditions. The minimum value is 0, this is when Tsol = 0 or ETC = 0 (pool full). We should set Tsol = 0 if Tsol < 0 (temperature below the freezing point: the water will remain at the soil surface). The maximum value is ETC, which can be reached, according to our examination of the data sets, by extremely high temperature (the observed extreme soil temperature was 62.7 degrees); we have set C = 70 (degrees).

An advantage of ETC is that it can be easily determined experimentally. In contrast to other variables, such as the fine fuel moisture, it is neither influenced by the stand structure, nor the topography and other sensitive factors; it represents the average dryness for the region, where it is measured or calculated.

2.3 Non-parametric estimate of the fire probabilities and assessment of the indexes

We use here the non-parametric estimate of the fire probabilities to assess the different indexes. (*Mandallaz* and *Ye,* 1996, *Ducharme et al.* 1995, see *Fig-*

ure 2). To this end we need a kernel function δ (.) and the correspondent cumulative kernel function K(.), e.g.:

Epachnikov's kernel: $\delta(x)=\frac{3}{4}\left(1-x^2\right)$, $\delta(x)=0$ *if* $x<-1$ *or* $x>-1$

$$K(x)=\frac{3}{4}\left(x-\frac{x^3}{3}+\frac{2}{3}\right),\ K(x)=0 \text{ if } x<-1 \text{ and } K(x)=1 \text{ if } x>-1$$

The fire probabilities at the index value z can be estimated according to:

$$\tilde{\pi}(z)=\frac{1}{Nb_2(N)}\sum_{j=1}^{N} I_j \delta\left(\frac{\tilde{F}_Z(z)-\tilde{F}_Z(z_j)}{b_2(N)}\right)$$

where $\tilde{F}_Z(z)=\frac{1}{N}\sum_{j=1}^{N} K\left(\frac{z-z_j}{b_1(N)}\right)$

N is the total number of days, $I_j=1$ if there was at least one fire on that day, otherwise $I_j=0$, z_j is the index value on the day j, $b_1(N)$ and $b_2(N)$ are the band width and should satisfy the following conditions:

$$\lim_{N\to\infty} b_k(N)=0$$

$$\lim_{N\to\infty} Nb_k(N)=\infty$$

The first condition (i) on the band width ensures asymptotic unbiasedness and consistency in mean square error of $\tilde{F}_Z(z)$, Condition (ii) essentially means that the number of data points within the bandwidth goes to infinity with N.

Theoretical and empirical studies suggest that the band width can be selected according to:

$$b_k(N)=\frac{c^*\,(\text{interquantile range of } z_j)}{N^{0.2}}$$

where c is a constant, which can be, according to our experience, selected between 0.5 and 2.0: the larger the c, the smoother the curve (and also the larger the bias, especially at both ends of the curve).

The main advantage of the non-parametric estimate of the fire probabilities is that it does not require the form of the relationship between the fire probabilities and the index values. However, a large number of observations is needed for a reliable estimation.

We say that an index is essentially correct if the true probability of fire on day i, π_i, is a monotone increasing function of the index value z_i of that day, that is $\pi_i=f(z_i)$. A good index should satisfy

1). $f(z)$ is monotone increasing
2). $f(z)$ as small as possible when z is near the minimum observed values (ideally zero)
3). $f(z)$ as large as possible when z is near the maximum observed values (ideally 1)

The form of the curves can be different for different indexes. *Figure 1* illustrates the curves of the non-parametric estimate of the fire probabilities of the three versions of our index. The x-axis is the index values and the y-axis is the estimated fire probability. Only data within the 90%-quantile of the curves are shown in the figure, because there are large bias where not enough observations are available.

According to *Figure 1*, ET is not bad for winter, except for the range of 0 to 0.4, where the fire probabilities are too high for the index values. This weakness becomes very clear in summer, where the fire probability at 0 is much higher than the average value. One reason for the high probabilities at lower index values is that ET is directly set to 0 on the rainy days. It is clearly not adequate under small precipitation, especially in summer.

In comparison with ET, ETC shows a clear improvement in this range, but the fire probabilities at the right part of the curve are now too low, and the disturbing peak-valley effect in the middle of the curve is enlarged.

The ETP curves illustrate the best characteristics among the three versions. The peak-valley has been «smoothed», the curves are essentially monotone increasing, the probabilities remain relatively low at index value 0, and high at the right part of the curves.

The peak-valley effect in the middle of the curves is partially due to the small number of observations, and also to the index itself. It can be improved, for instance, by replacing the evaporation of water with the evapotranspiration of the forest (see also *Mandallaz* and *Ye,* 1996), but the improvement is negligible in comparison with the increase of the complexity of the index.

Figure 2 illustrates several indexes, calculated with the same data sets used in this study. The Poisson models give the predicted fire probabilities as an index, by incorporating various variables into the models, such as the dryness index, indicator variables describing the geographical difference, fire seasons, social-economical factors, and the short term fire history (*Mandallaz* and *Ye,* 1997).

The curves of the Poisson models are the best we have found. They begin with a fire probability of 0 and go very smoothly up to a relatively high fire probability (in comparison with other indexes). This quality cannot be reached by using a dryness index alone. Among the single indexes, the curve of IFM is very good in winter, and also acceptable in summer. But the curves of IREPI and RN are not so good, especially in the lower range of the index values.

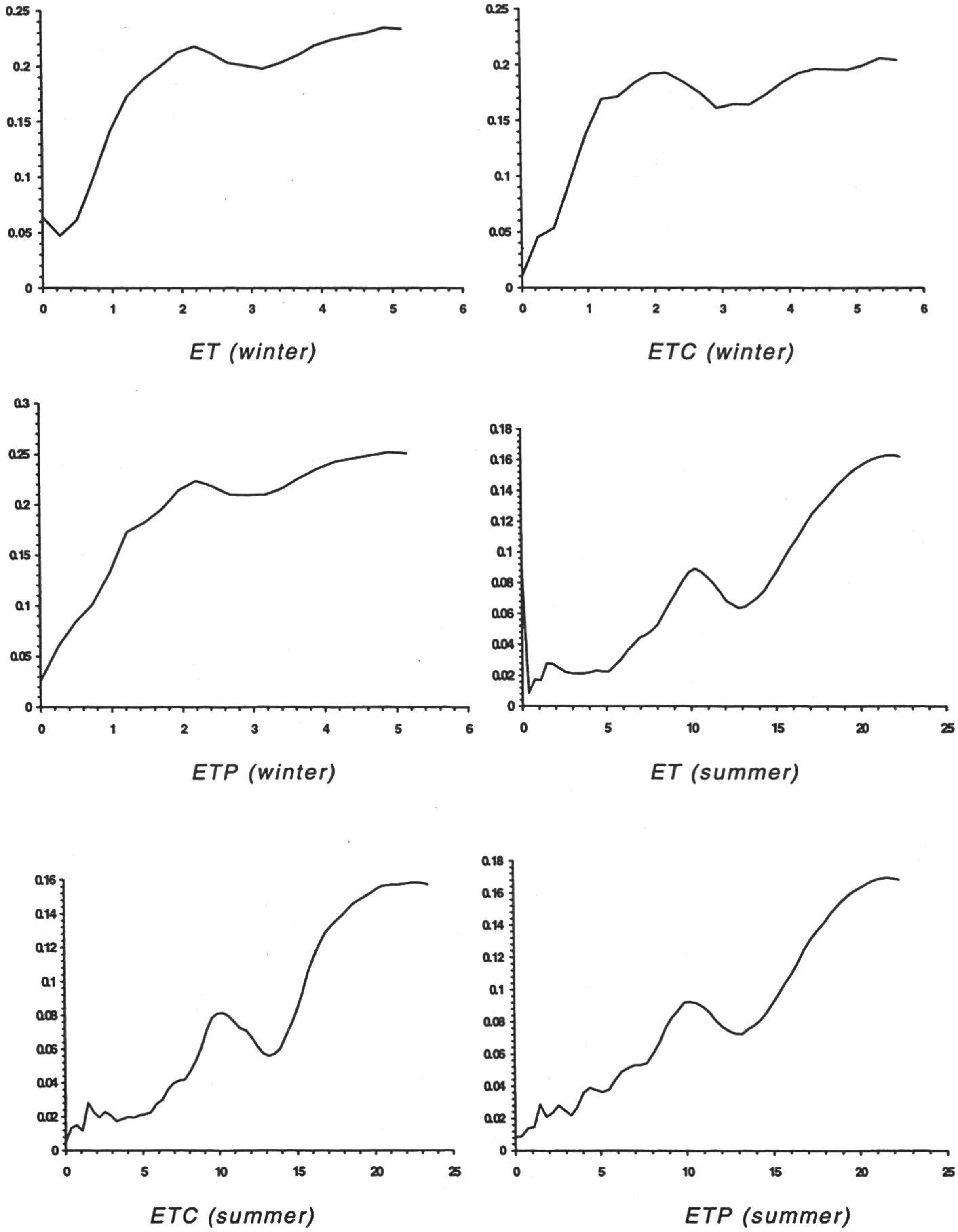

Figure 1. Non-parametric estimate of the fire probabilities for ET, ETC and ETP.

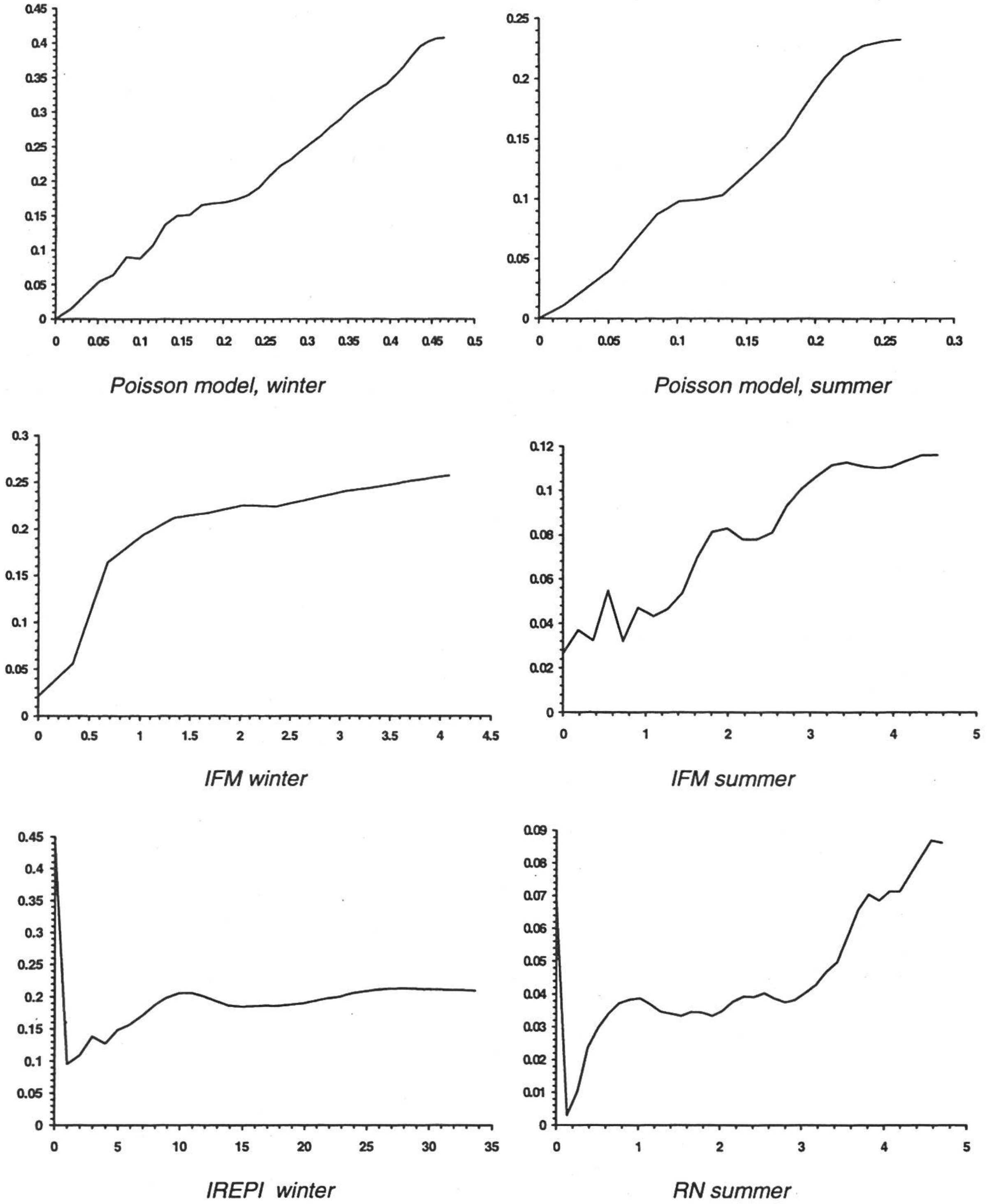

Figure 2. Non-parametric estimate of the fire probabilities for different indexes.

3. Comparison of the performance of different indexes

In conjunction with the curves, other methods can also be used for the assessment of indexes. We use here three methods suggested by *Mandallaz* and *Ye* (1996).

The first criterion is the number of correctly predicted fire days under a selected decision rule (for a discussion of different decision rules see *Mandallaz* and *Ye,* 1996). According to a decision rule, threshold values can be defined for each index. If the index value of a day is larger than the corresponding threshold, that day will be predicted as a «fire day». If the same number of fire days is to be predicted with every method, the number of correctly predicted fire days can be used to compare the «prediction ability» of each method.

We use here the decision rule ensuring that the number of predicted fire days is equal to the number of observed fire days for the whole data set (from 1980 to 1995), a very reasonable requirement. The results are listed in *Table 1.*

According to this criterion, the Poisson models have the best performance in comparison with all pure meteorological indexes. The best indexes are RN, IFM, IP, ETP for winter, and IREPI, IP, ETP for summer. For the main fire season the differences among the meteorological indexes are small (21–27% correct prediction).

Table 1. Prediction ability of the different methods.

Methods	*Number of correctly predicted fire days*			
	winter	*correct %*	*summer*	*correct %*
Random decision*	65.58	9.5	17.17	4.1
Index IREPI	146	21	85	20
Index ETP	156	23	74	18
Index RN	187	27	37	9
Index IFM	176	26	56	13
Index ICONA	146	21	57	14
Index IP	176	26	80	19
Poisson model	267	39	123	29
Total days / fire days	7260/690		10272/420	

* the number of correctly predicted fire days for random decision is calculated as p2*N, where p is the observed fire probability.

It can be noted that the index RN (winter) and IREPI (summer) have more correctly predicted fire days than other meteorological indexes according to this decision rule, even though their fire probability curves are not better. We have found out that this is not the case when using other decision rules. For «essentially correct» indexes the result of the comparison will be consistent according to different decision rules. This criterion alone is, therefore, not suitable for the assessment of the overall performance of the indexes.

The second criterion, which can be used for the assessment of the overall performance of the indexes, is a score, defined as:

$$S = (I - I_{random}) / (I_{max} - I_{random}),$$

where: $I = \sum_{i=1}^{N} rank(z_i) I_i$ N is the number of total days, z is an index, $rank(.)$ the rank of the index value within the data set, $I_i = 1$ if there was at least one fire on that day, otherwise $I_i = 0$.

We examine 2 special cases: 1) the deterministic case. All the d fire days correspond to the largest d index values and all of them are correctly predicted. In this case I has the maximum value $I_{max} = d(2N+1-d)/2$ and the score S is 1. 2) the random case. The index z_i is stochastically independent of I_i, the average rank of the index is $(N+1)/2$ and the expected mean value of I is then $I_{random} = d(N+1)/2$ and the score S is 0. The other indexes, should have the score value between *0* and *1* (provided they are better than the random index). The results are listed in *table 2.*

Table 2. Performance score S of the different methods.

Methods	*Score*	
	winter	*summer*
Index RN	0.46	0.38
IREPI	0.43	0.39
Index ETP	0.53	0.55
ICONA	0.42	0.43
Index IP	0.51	0.50
Index IFM	0.58	0.43
Poisson model	0.66	0.66

According to *table 2*, the best methods are: Poisson models, Index IFM, Index ETP for winter, and Poisson models, Index ETP, Index IP for summer.

The third criterion we propose is a qualitative tool, which provides a visual assessment of the overall performance of the indexes (see *Figure 3*). *Figure 3* gives the so called cumulative distribution of the fire frequencies. They are defined as:

$$C_N(x) = C_N\left(\frac{i}{N}\right), \frac{i}{N} \leq x < \frac{i+1}{N}$$

$$C_N\left(\frac{i}{N}\right) := \frac{1}{N}\sum_{j=1}^{i} I_{\{rank(z_j) \leq i\}} I_j \ , i = 1, 2, \ldots, N$$

where $I_{\{rank(z_j) \leq i\}}$ = 1, if the rank of index value z_j in the whole data set is smaller than or equal to i, otherwise 0.

For the deterministic case the curve is a broken line (MAX in *Figure 3*). For the random case the curve is a straight line (RANDOM in *Figure 3*). The other curves lie in-between (which have been smoothed according to *Ducharme et al.* 1995).

Mandallaz and *Ye* (1997) have shown that the curves of essentially correct indexes must be convex. Obviously, they should also be as far away as possible from the random line, and as close as possible to the deterministic broken line.

In *Figure 3*, the x-axis is the rank of the index values, scaled to the range [0,1]. The y-axis is the cumulative fire frequencies, which begin at 0 and reach the average observed fire probability for the data set. The small biases at the end points are due to the smoothing techniques.

Whereas the Poisson models always show the best properties, the other indexes have very different performances. In winter the IFM and ETP curves are close together and nicely convex, though ETP is not as good as IFM in the middle range. In summer ETP is clearly better than the other indexes, the next best is IP. The curves of the indexes ICONA, IP, IREPI, RN for winter as well as IREPI for summer are not convex, and cannot be considered as essentially correct.

Summarising the overall performance of the different methods according to the score, the cumulative distribution of the fire frequencies and the curves of the non-parametric estimate of the fire probabilities, we can state that a good index will have good performances according to these criteria. The consistence of the three criteria can be observed in other data sets, from France, Italy and Portugal (see *Mandallaz* and *Ye,* 1996). However, the three criteria reveal different aspects of the indexes, which should all be taken into account.

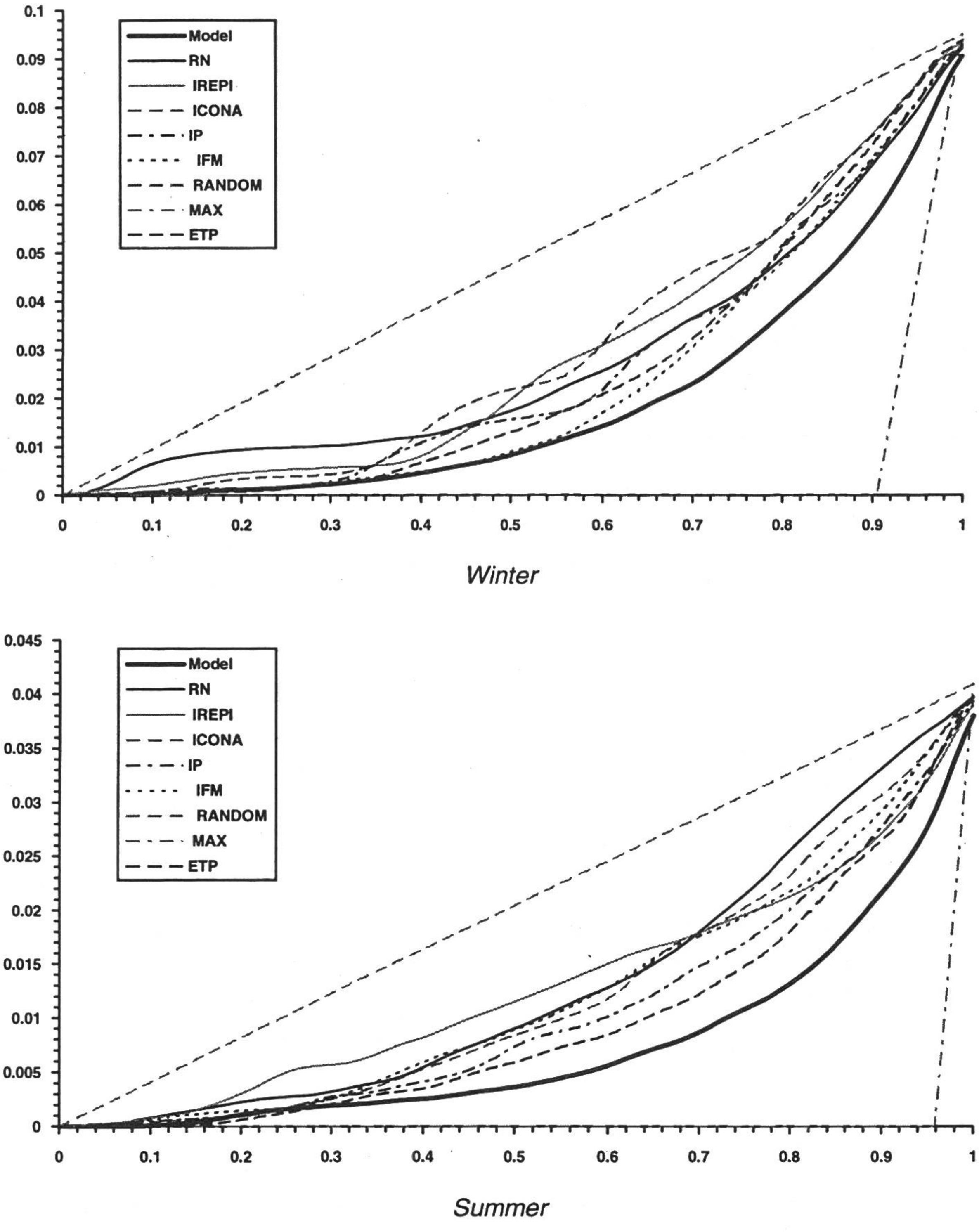

Figure 3. Cumulative distribution of the fire frequencies vs. scaled index values.

4. Conclusions

Based on extensive case studies the following conclusions can be drawn:

1. The index ETP is very simple, has a very clear physical interpretation. It is easy to calculate or measure, and has good performance.
2. The non-parametric estimate of the fire probabilities is a useful tool to detect spurious behaviour of an index; it is reliable for large data sets.
3. The quantitative and qualitative assessment of indexes should be based on the score, the cumulative distribution of the fire frequencies, and the behaviour of the non-parametric estimate of the fire probabilities.
4. For the southern Alps of Switzerland the best meteorological indexes are IFM and ETP for winter and ETP and IP for summer.
5. Pure meteorological indexes have poorer performance than indexes based on the direct modelization of fire possibilities by means of Poisson models incorporating further explanatory variables.

Summary

A new dryness index for the fire danger prediction is presented in this paper. This index is intuitively very simple, easy to implement in practice and has a high performance in comparison with other indexes used.

In order to find out the weakness of an index the non-parametric estimate of the fire probabilities based on the index is used; this is particularly useful to avoid large conceptual mistakes in the construction of the index. Furthermore, it is possible to give a simple criterion for an index to be essentially correct. Comparison studies have shown that indexes satisfying this criterion will have higher overall performance than those which do not.

Zusammenfassung

Ein neuer Trockenheitsindex und die nicht-parametrische Schätzung der Waldbrandwahrscheinlichkeit

In diesem Aufsatz wird ein neuer einfacher Trockenheitsindex für die Waldbrandvorhersage vorgestellt. In Beispielen wird gezeigt, wie die Stärken und Schwächen der Indexe mit den Schätzungen der Brandwahrscheinlichkeit beurteilt werden können. Die Güte von verschiedenen Indexen wird verglichen. Gestützt auf die meteorologischen und die Waldbranddaten aus der Schweizer Alpensüdseite von 1980 bis 1995 können die folgenden Schlussfolgerungen gezogen werden:

1. Der neue Index ETP hat eine klare physikalische Bedeutung und ist einfach zu berechnen oder zu messen. Zudem weist er eine erstaunlich hohe Güte auf.
2. Die nicht-parametrische Schätzung der Waldbrandwahrscheinlichkeit ist ein nützliches Werkzeug für die Beurteilung eines Trockenheitsindexes. Sie braucht jedoch

eine relative hohe Anzahl von Beobachtungen, um eine zuverlässige Schätzung zu machen.

3. Die Beurteilung der Güte eines Indexes auf Grund der Schätzung der Brandwahrscheinlichkeit, der kumulativen Verteilung der Waldbrandhäufigkeit und verschiedener Bewertungskriterien zeigt die verschiedenen Aspekte eines Indexes auf. Ein guter Index sollte nach allen drei Kriterien gut abschneiden.
4. Die besten rein meteorologischen Waldbrandgefahrindexe waren für die Alpensüdseite IFM und ETP für Winter und ETP und IP für Sommer.
5. Die mittels Poisson-Modellen abgeleiteten multifaktoriellen Gefahrindexe waren mit Abstand besser als die rein meteorologischen Indexe.

Résumé

Un nouvel indice de sécheresse et l'estimation non-paramétrique de la probabilité d'incendies

Cet article présente un nouvel indice de sécheresse pour la prévision des incendies de forêts. Des exemples démontrent que l'estimation non-paramétrique de la probabilité d'incendies est un outil précieux pour déceler les faiblesses et qualités d'un indice. Sur la base des incendies survenus dans la période 1980–1995 dans le Sud des Alpes suisses l'on peut tirer les conclusions suivantes:

1. Le nouvel indice ETP possède une interprétation physique évidente et est particulièrement facile à calculer ou mesurer.
2. L'estimation non-paramétrique de la probabilité d'incendies, la distribution cumulative de l'indice de danger ainsi que le taux de prédictions correctes sont 3 critères permettant d'évaluer les différents aspects d'un indice. Un bon indice remplit ces 3 critères.
3. Les indices de danger basés sur la modélisation de la probabilité d'incendie à l'aide de modèles de Poisson multifactoriels sont toujours supérieurs aux indices purement météorologiques.

Literature

Bovio G., Sol B. and *Viegas D.X.* 1994. Studies synthesis on the inter-comparison of meteorological fire hazard indexes. Programme Minerve 1, European Economic Community (DG XII).

Ducharme G. R. et al. 1995. Reference values obtained by kernel-based estimation of quantile regression, Biometrics, *51*:1105–1116.

Item, H., 1974. Ein Modell für den Wasserhaushalt eines Laubwaldes, Mitteilungen der Eidg. Anstalt für das Forstl. Versuchswesen. *50*:3, 137–331.

Item, H., 1981. Ein Wasserhaushaltmodell für Wald und Wiese, Mitteilungen der Eidg. Anstalt für das Forstl. Versuchswesen. *57*:1, 5–82.

Mandallaz, D. and *Ye, R.* 1996. Statistical Model for the prediction of forest fires, Final report MINERVE II, European Economic Community (DG XII).

Mandallaz, D. and *Ye, R.* 1997. Prediction of forest fires with Poisson models, Can. J. For. Res., (accepted for publication).

Marcozzi M., Bovio G., Mandallaz D. and *Bachmann P.* 1994. Influenza della meteorologia sull'indice di pericolo degli incendi boschivi nel Cantone Ticino, Schweiz. Z. Forstwes. *145*, 3:183–199.

Sol, B. 1995. Comparaison de diverses méthodes d'estimation du danger météorologique d'incendie sur le sud-est, METEO-FRANCE, Note DIR/SE N0. 13.

Acknowledgement

We express our thanks to P. Ambrosetti (Swiss Meteo-Service) and M. Conedera (WSL) for providing the data sets, and to the European Community and the Swiss Ministry of Science and Education for the financial support.

Appendix: SAS program for the calculation of the indexes ET, ETC and ETP

```
data etetp;
  set meteo;
  retain ETC 0;                  /* start value for ET, ETC and ETP */
  retain rainc 0 ett 0;          /* rainc for rest of precipitation */
                               /* ett is a temporal variable for ETP*/
  /* Penman's formula, Evaporation ep*/
  Iq = (1E7*Insolg/2.778)*0.94;     /* 0.94 is the Albedo for water */
  n_N = insolr/100;                 /* relative sun radiation       */
  Kair = Tair + 273;                /* absolute temperature         */
  Ea = 6.1121E-3*exp(17.67*Tair/(Tair+243.5));
  Ed = (Hum/100)*Ea;
  L = 3143.58-2.36*Kair;
  Delta = L*Ea/(0.461*Kair*Kair);
  B = 4.90011*(Kair**4)*(0.56-(2.47*sqrt(Ed)))*(0.1+0.9*n_N);
  R = Iq - B;
  gamma = 0.6687E-3;
  Ep = 0.4E-10*Delta*R + (100*gamma*(0.131 + 0.0014*v1)*(Ea-Ed));
  Ep = Ep/(Delta + gamma);
  if Ep <= 0 and Ep ne . then Ep = 0;
  /* Penman's formula, end */

 ETC + (Ep-rain);                  /* rain is precipitation, in cm  */
 if ETC < 0 then ETC = 0;          /* overflow of the pool          */
 ET = ETC;
 if Ep-rain < 0 then ET = 0;       /* raining day for ET            */

 if ep<rain then do;                   /*ETP: on the day of raining */
    tt=tsol*etc/70;
    if tt<0 then tt=0;
    ett=tt;
    rainc=(rain-ep);
    etp=ett;
 end;
 else do;                                       /*  following days */
   if ep-rain >=rainc then do;              /*without rest rainfall */
    etp=etc;
    rainc=0;
    ett=0;
   end;
   else do;                                  /*  with rest rainfall */
    tt=tsol*etc/70;
    if tt<0 then tt=0;
    ett+tt;
    rainc+(rain-ep);
    etp=ett;
   end;
 end;

 if etp>etc then etp=etc;
 keep datum region Nfire ET ETC ETP;
 run;
```


Authors:
PD Dr. Daniel Mandallaz, Chair of Forest Inventory and Planning, ETH Zurich, CH-8092 Zurich, Switzerland; e-mail: mandallaz@waho.ethz.ch.
Dr. Ronghua Ye, Academy of Forest Inventory and Planning, Hepingli Dongjie 18, 100714 Beijing, China.