Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

EGU General Assembly 2020

cen

Global (and regional) performances of SPI candidate distribution functions in observations and simulations

Patrick Pieper(1), André Düsterhus(2), and Johanna Baehr(1)

06th May 2020

Sharing Geoscience Online: HS3.6:
Spatio-temporal and/or (geo) statistical analysis of hydrological events, floods, extremes, and related hazards

1 – Institute for Oceanography, Center for Earth System Research and Sustainability, Universität Hamburg, Hamburg, Germany

2 – ICARUS, Department of Geography, Maynooth University, Maynooth, Ireland

Quick-Story

Aim:

Find a single PDF that ensures SPI‘s standardization universally for all:
locations of the globe during all accumulation periods and in observations + simulations.

Method:

SPI time-series derived with:
- Gen. gamma distribution (GGD3)
- Weibull distribution (WD2)
- Exp. Weibull distribution (EWD3)
- Gamma distribution (GD2)

What we analyze

SPI time-series derived from:
- Max-Planck-Institute Earth System Model
- Global Precipitation Climatology Project

How we analyze

- Comparing actual against theoretical occurrence probability (TOP)
- Ranking by Bayesian Information Criterion Differences (BIC-D)

TOP and BIC-D form complementary analysis:
- TOP relies on a subjective evaluation, BIC-D performs analytically evaluation.
- BIC-D assesses relative performance differences, TOP assesses performance in absolute terms.

Result:

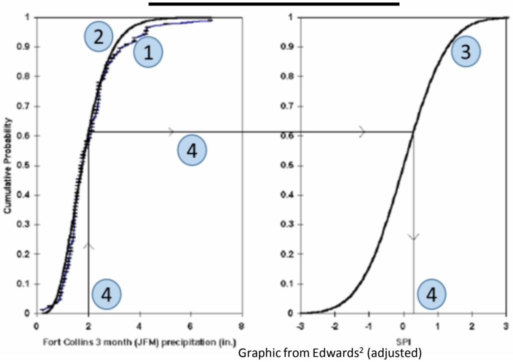
Identification of a single PDF that universally ensures SPI’s standardization:
the 3-parameter exponentiated Weibull distribution.

Conclusion:

Flawless performance of exp. Weibull distribution across all targeted dimensions:
unprecedented in SPI-literature.

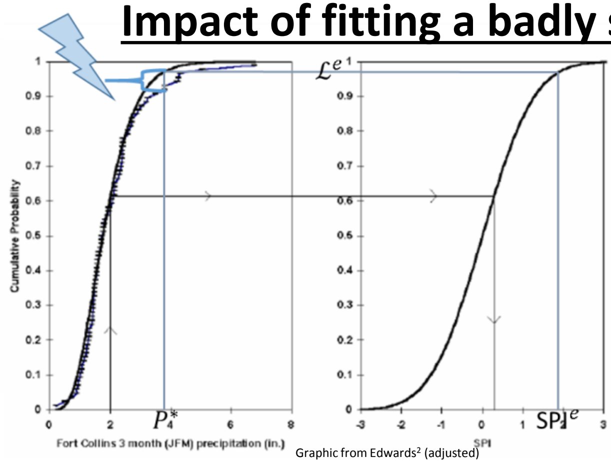
Background: Standardized Precipitation Index (SPI)

Calculation[1]:



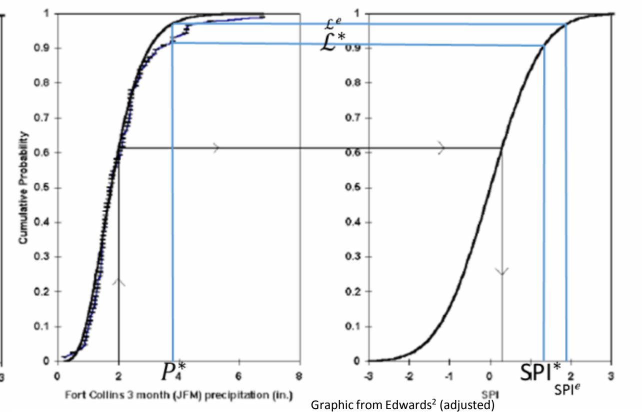
1. Sort precip. totals for chosen accumulation period (displayed here as empirical cumulative PDF).
2. Fit candidate PDF onto sorted precip. totals. (Fit performed onto modeled precip. for all ensemble members at once.)
3. Z-transformation of cumulative probabilities onto standard normal PDF (μ=0; σ=1).
4. Assign precipitation totals of initial time-series to corresponding SPI values.

Impact of fitting a badly suited candidate PDF:



Encountering precip. totals of $\leq P^*$ is erroneously assigned the likelihood $\mathcal{L}^e$. That results in the erroneous SPI value SPI^e.

Selecting a suitable candidate PDF when calculating SPI is of paramount importance to preserve SPI's standardization.



In contrast, the likelihood $\mathcal{L}^*$ is a more realistic estimate of the occurrence probability of a precip. value of $\leq P^*$. I.e. the fit should be closer to $\mathcal{L}^*$ at P^*.

What's the big deal?

$\mathcal{L}^*$ is roughly 3 times less likely than $\mathcal{L}^e$ (transgressing the ~97th percentile vs. the ~91st percentile).

Thus, the resulting SPI time-series would displays roughly 3 times as many SPI values $\geq$ SPI^e than expected from a standard normally distributed variable.

Motivation

Advantages[3,4]: **Standardization, spatio-temporal comparability, simplicity**, …

Disadvantage[4]: Realization of its standardization: Thread to undermine each advantage.

- Key decision: Choice of suitable candidate probability density function (PDF)[3,5,6].
- Obstacle: That PDF should describe highly non-normal precipitation distributions[4].
- Preserving advantages aggravates obstacle:
  Surmount obstacle concurrently ∀:
  - Accumulation periods.
  - Locations around the globe.
  - In observations and **simulations (gap in lit.)**.

Insurmountable by 2-parameter PDFs[3-10].
Thus, when using SPI one should first:
test normality anew for investigated data sets, locations, and accumulation period[3-8].

→ Repetitive normality tests for candidate PDFs undermine SPI's advantage of **simplicity**.
→ Use of different PDFs undermine SPI's advantage of **spatio-temporal comparability**.
→ Simply employing one PDF irrespective of universal applicability violates **standardization**.

Obstacle seems too complex for a quick-fix.
Particularly when considering and tackling the gap in literature (simulations) concurrently.

→ Investigate more complex solutions: 3-parameter PDFs.

Methods: Model and Data

Modeled Precipitation: Max-Planck-Institute Earth System Model (MPI-ESM):

- Seasonal Prediction System[11]: Initialized in May and November with 10 ensemble members.
- Run in Low-Resolution (MPI-ESM-LR): T63 with 47 (40) layers in the atmosphere (ocean).
- Neither bias- nor drift-corrected: Analyze PDF's fit of modeled distribution → worst-case.

Observed Precipitation: Global Precipitation Climatology Project (GPCP)[12]:

- Combines direct observations and satellite-data.
- Monthly precipitation set on a 2.5°x2.5° global grid.

Ensuring spatio-temporal Comparability between Model and Observations:

- Modeled precip. (on T63 – approx. 1.875°x1.875°) interpolated to GPCP's grid (**2.5°x2.5°**).
- Analyze common time period of hindcasts (**1982-2013**) and GPCP (1979-present).

Global and regional investigation:

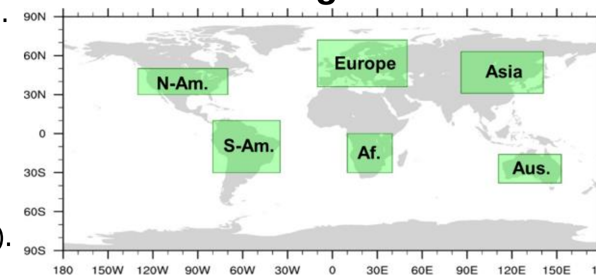


***Figure 1**: Borders of analyzed regions*

Methods: Candidate PDFs

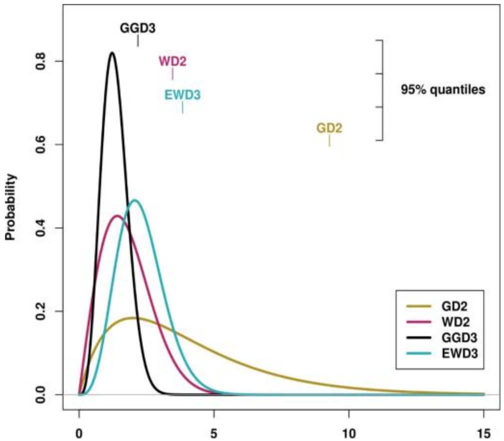


***Figure 2**: Analyzed candidate PDFs. Displayed are examples of those PDFs with scale and shape parameter(s) = 2.*

Analyzed candidate PDFs:

- **Gamma distribution (GD2)**
- **Weibull distribution (WD2)**
- **Gen. gamma distribution (GGD3)**
- **Exp. Weibull distribution (EWD3)**

Distribution function	Parameter count	Abbreviation
Gamma distribution	2	GD2
Weibull distribution	2	WD2
Generalized gamma distribution	3	GGD3
Exponentiated Weibull distribution	3	EWD3

Fitting Procedure:

Parameter estimation method:

- Maximum-Likelihood Estimation (MLE)

Optimization procedure for MLE:

- Simulated annealing method[13]
- BFGS quasi-Newton method[14]
- Nelder and Mead method[15]

Methods: Skill Metrics

1. Comparing actual against theoretical occurrence probability (TOP)[3]:

- Needs definition of analyzed SPI intervals. → 7 different classes are usually discussed in literature[16]:

SPI intervals	SPI≥2	2>SPI≥1.5	1.5>SPI≥1	1>SPI>-1	-1≥SPI>-1.5	-1.5≥SPI>-2	SPI≤2
SPI classes	Extremely wet	Severely wet	Moderately wet	Normal	Moderately dry	Severely dry	Extremely dry
TOP [%]	2.3%	4.4%	9.2%	68.2%	9.2%	4.4%	2.3%

2. Ranking by Bayesian Information Criterion Differences (BIC-D)[17]:

Bayesian Information Criterion (BIC) analytically evaluates trade-off: information gain vs. complexity (punishes complexity).
BIC-D discriminate PDFs based on relative difference to best-performing PDF: $\text{BIC−D}_i = \text{BIC}_i - \text{BIC}_{min}$
while i indexes candidate PDFs, min indicates the best performing PDF.

- Cannot evaluate absolute performance
  → Needs to be safeguarded by other metric(s)!
- Interpretation according to Burnham and Anderson[17]:

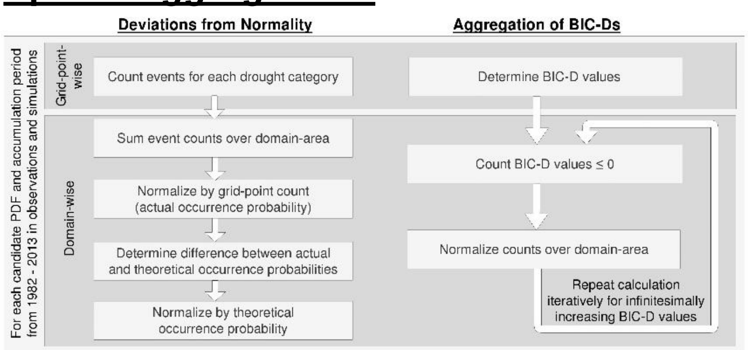
BIC-D value	< 2	< 4	< 7	> 10
BIC-D interpretation	ideal	well	sufficient	no skill

TOP and BIC-D form complementary analysis:

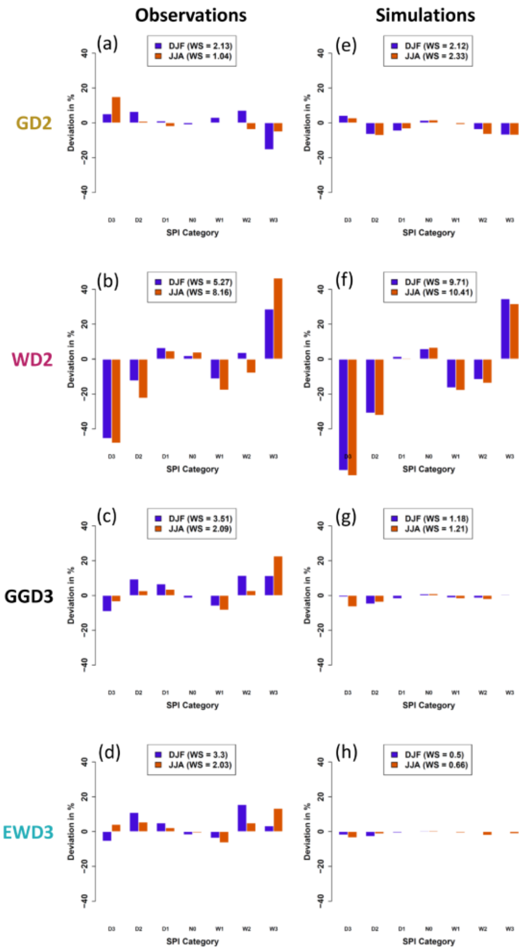
- While TOP comparison (1) relies on a subjective evaluation, BIC-D rank (2) PDF based on an analytically evaluation.
- While BIC-D only assess relative performance differences, our TOP comparison assess performance in absolute terms.

	TOP	BIC-D
Comparison	Subjective	**Analytical**
Assessment	**Absolute**	Relative

Spatial Aggregation:



Results: SPI$_{3M}$ Global deviations from TOP



Deviations from the Normal Distribution:

- **GD2**, **GGD3**, and **EWD3** describe without seasonality similarly well the overall frequency distribution of observed 3-months precipitation totals.
- **WD2** performs overall poorly and is in every regard inferior to any other candidate PDF.
- **GGD3** and **EWD3** describe the frequency distribution of modeled 3-months precipitation totals distinctly better than any 2-parameter candidate PDF.
- **GD2** describes the frequency distribution of modeled 3-months precipitation totals still sufficiently well on global average.
- Both 2-parameter candidate PDFs are unable to benefit from the increased length of the database in simulations (fit all ensemble members at once) relative to observations, while both 3-parameter PDFs strongly benefit from that increase. Apparent in weighted (by TOP) sum of deviations (WS in legend).

No seasonal-dependence of absolute performance:
Without flaws for best-suited PDFs (**EWD3**, **GGD3**, and **GD2**).

Open question: performance of **EWD3/GGD3** relative to **GD2:**
Enough **improvement** to justify increased **complexity**?

***Figure 3**: Deviations between actual and TOP. Displayed for observed (left) and modeled (right) SPI time-series. SPI time-series are derived by using the simple 2-parameter gamma distribution (GD2, top row), the simple 2-parameter Weibull distribution (WD2, second row), the 3-parameter generalized gamma distribution (GGD3, third row), and the 3-parameter exponentiated Weibull distribution (EWD3, bottom row). The legends depict the sum of deviations along all SPI categories weighted by their respective theoretical occurrence probability: weighted sum (WS).*

Results: SPI_{3M} Global BIC-D

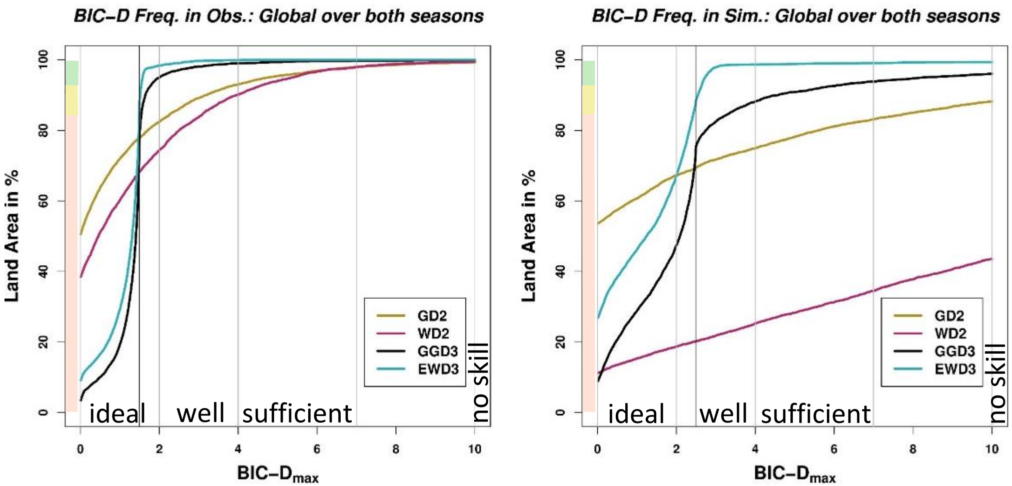


***Figure 4**: BIC-D frequencies. Percentages of global land grid-points in which each distribution function yields BIC-D values that are smaller than or equal to a given BIC-D_{max} value. BIC-D freuqencies sum up to 100% at the BIC-D_{max} value of 0 (only one PDF performs best in each grid-point). Vertical black lines indicate increased complexity penalty (CP) of 3- relative to 2-parameter PDFs (sample size-dependent). BIC-D frequencies are displayed for each candidate PDF for observations (left) and simulations (right).*

***Table**: Percent of grid-points which are classified according to Burnham and Anderson depending on whether they display BIC-D values lower than specific thresholds or higher than 10 for each candidate PDF over both seasons. Percentages of grid-points indicate the confidence in candidate PDFs to overall perform according to the respective BIC-D category.*

SPI Period	Realization	BIC-D category	GD2	WD2	GGD3	EWD3
3-Months	Observations	Ideal (BIC-D $\leq$ 2)	82	74	95	98
		Well (BIC-D $\leq$ 4)	93	90	99	100
		Sufficient (BIC-D $\leq$ 7)	98	98	100	100
		No Skill (BIC-D > 10)	1	1	0	0
	Simulations	Ideal (BIC-D $\leq$ 2)	67	19	47	67
		Well (BIC-D $\leq$ 4)	75	25	88	99
		Sufficient (BIC-D $\leq$ 7)	83	34	94	99
		No Skill (BIC-D > 10)	12	56	4	1

BIC-D Frequencies:

Percentages of grid-points indicate the confidence in candidate PDFs to overall perform according to the respective BIC-D category. We evaluate coverages as follows:

≥ 95% (≤ 5%) as sign of substantial confidence.

≥ 85% (≤ 15%) as sign of average confidence.

< 85% (> 15%) as sign of insufficient confidence.

Performances of candidate PDFs in Observations:

- **EWD3** ideal in virtually every grid-point.
  - Without any flaws.
- **GD2** flawed in ~7% of grid-points (GP).

Simulations:

- **EWD3** ideal in as many GPs as GD2.
  - Without any flaws.
- **GD2** ideal in in as many GPs as EWD3.
  - But flawed in ~25 % of grid-points.
  - And worthless in ~12% of GPs.

Summary

- Complementary test methodology for SPI normality:
  - Difference between actual and TOP as interest lies on classes with well-defined intervals.
  - Alongside BIC-D to accompany subjective evaluation by an analytical, user-, and aggregation-friendly metric. BIC-D should not be used as stand-alone metric!

	TOP	BIC-D
Comparison	Subjective	**Analytical**
Assessment	**Absolute**	Relative

- **EWD3** – **SPI's best-suited candidate PDF** in our analysis: Standardization ∀:
  - ☑ Accumulation periods (not shown, here).
  - ☑ Locations around the globe (not shown. But global results are robust in each region).
  - ☑ In observations and simulations.

SPI Period	Realization	BIC-D category	GD2	WD2	GGD3	EWD3
3-Months	Observations	Ideal (BIC-D $\leq$ 2)	82	74	95	98
		Well (BIC-D $\leq$ 4)	93	90	99	100
		Sufficient (BIC-D $\leq$ 7)	98	98	100	100
		No Skill (BIC-D $>$ 10)	1	1	0	0
	Simulations	Ideal (BIC-D $\leq$ 2)	67	19	47	67
		Well (BIC-D $\leq$ 4)	75	25	88	99
		Sufficient (BIC-D $\leq$ 7)	83	34	94	99
		No Skill (BIC-D $>$ 10)	12	56	4	1

Full-Story

- Goodness-of-Fit tests cannot analyze SPI-normality.
- Performances of candidate PDFs for all common accumulation periods (1-, 6-, 9- , and 12-months).
- **EWD3** is **also better-suited** than (indistinguishable from) a **multi-PDF** approach.
- The meticulousness applied to the optimization procedure is as important as the choice of the candidate PDF (and probably more important than the parameter estimation method).
- **EWD3** outperforms the commonly employed **GD2** for a sample size of 31 years.
  The larger the sample size, the larger is also the improved performance.

Full-story available in HESSD:

Pieper, P., Düsterhus, A., and Baehr, J.: Global and regional performances of SPI candidate distribution functions in observations and simulations, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-614, in review, 2020.

References I

1. McKee, T. B. et al.: The relationship of drought frequency and duration to time scales, in: Proceedings of the 8th Conference on Applied Climatology, vol. 17, pp. 179–183, American Meteorological Society Boston, MA, 1993.
2. Edwards, D.C. and McKee, T. B.: Characteristics of 20th Century Drought in the United States at Multiple Scales, Atmospheric Science Paper No. 634, May 1–30, 1997.
3. Sienz, F., Bothe, O., and Fraedrich, K.: Monitoring and quantifying future climate projections of dryness and wetness extremes: SPI bias, Hydrology and Earth System Sciences, 16, 2143, 2012.
4. Lloyd-Hughes, B. and Saunders, M. A.: A drought climatology for Europe, International Journal of Climatology: A Journal of the Royal Meteorological Society, 22, 1571–1592, 2002.
5. Blain, G. C. and Meschiatti, M. C.: Inadequacy of the gamma distribution to calculate the Standardized Precipitation Index, Revista Brasileira de Engenharia Agrícola e Ambiental, 19, 1129–1135, 2015.
6. Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Van Loon, A. F., and Stahl, K.: Candidate distributions for climatological drought indices (SPI and SPEI), International Journal of Climatology, 35, 4027–4040, 2015.
7. Guenang, G., Komkoua, M., Pokam, M., Tanessong, R., Tchakoutio, S., Vondou, A., Tamoffo, A., Djiotang, L., Yepdo, Z., and Mkankam, K.: Sensitivity of SPI to Distribution Functions and Correlation Between its Values at Different Time Scales in Central Africa, Earth Systems and Environment, pp. 1–12, 2019.
8. Touma, D., Ashfaq, M., Nayak, M. A., Kao, S.-C., and Diffenbaugh, N. S.: A multi-model and multi-index evaluation of drought characteristics in the 21st century, Journal of Hydrology, 526, 196–207, 2015.
9. Blain, G. C., de Avila, A. M. H., and Pereira, V. R.: Using the normality assumption to calculate probability-based standardized drought indices: selection criteria with emphases on typical events, International Journal of Climatology, 38, e418–e436, 2018.
10. Naresh Kumar, M., Murthy, C., Sesha Sai, M., and Roy, P.: On the use of Standardized Precipitation Index (SPI) for drought intensity assessment, Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 16, 381–389, 2009.

References II

11. Baehr, J., Fröhlich, K., Botzet, M., Domeisen, D. I., Kornblueh, L., Notz, D., Piontek, R., Pohlmann, H., Tietsche, S., and Mueller, W. A.: The prediction of surface temperature in the new seasonal prediction system based on the MPI-ESM coupled climate model, Climate Dynamics, 44, 2723–2735, 2015.
12. Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., et al.: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present), Journal of hydrometeorology, 4, 1147–1167, 2003.
13. Bélisle, C. J.: Convergence theorems for a class of simulated annealing algorithms on Rd, Journal of Applied Probability, 29, 885–895, 1992.
14. Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C.: A limited memory algorithm for bound constrained optimization, SIAM Journal on Scientific Computing, 16, 1190–1208, 1995.
15. Nelder, J. A. and Mead, R.: A simplex method for function minimization, The computer journal, 7, 308–313, 1965.
16. Svoboda, M., Hayes, M., and Wood, D.: Standardized precipitation index user guide, World Meteorological Organization Geneva, Switzerland, 2012.
17. Burnham, K. P. and Anderson, D. R.: Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach, 2002.