

## تاثیر عوامل ورودی بر احتمال‌های پاسخ از مرتبه‌ی $p$

فرزاد اسکندری<sup>۱</sup>، محمد خراسانی<sup>۲\*</sup>

<sup>۱</sup> دانشکده آمار، ریاضی و رایانه، دانشگاه علامه طباطبائی، تهران، ایران  
<sup>۲</sup> دانشکده آمار، ریاضی و رایانه، دانشگاه علامه طباطبائی، تهران، ایران

mokhorasani555@gmail.com

**چکیده:** در دنیای امروزی با استفاده از فرآیند مدل‌سازی آماری می‌توان پدیده‌های طبیعی را جهت تحلیل و پیش‌بینی اتفاقات مورد بررسی، بکار برد. به‌طور مشخص در هنگام پیش‌بینی، پارامترها بر عملکرد مدل‌های آماری تاثیر می‌گذارند. در بسیاری از مسائل ارزیابی ریسک، وجود عدم قطعیت در پارامترها منجر به عدم قطعیت در پیش‌بینی‌های مدل می‌شود. تحلیل حساسیت کلی ( $GSA$ ) ابزاری است که برای نشان‌دادن عدم قطعیت استفاده می‌شود و در تصمیم‌گیری، ارزیابی ریسک، ساده‌سازی مدل‌ها و غیره کاربرد دارد. تحلیل حساسیت فاصله مینکوفسکی و تحلیل حساسیت منطقه‌ای دو روش پرکاربرد هستند که می‌توانند با یک مجموعه نمونه داده‌شده از جفت‌های ورودی - خروجی مدل کار کنند. یک تفاوت قابل توجه بین این دو روش آن است که تحلیل حساسیت فاصله مینکوفسکی، توزیع‌های خروجی را مشروط به مقادیر ورودی (پیشرو) تحلیل می‌کند در حالی که تحلیل حساسیت منطقه‌ای، توزیع‌های ورودی را مشروط به مقادیر خروجی تحلیل می‌کند (معکوس). در این مقاله به تعیین روابط روش‌های حساسیت کلی (فاصله مینکوفسکی و منطقه‌ای) پرداخته و نشان داده خواهد شد، هنگامی که تحلیل حساسیت منطقه‌ای بر توابع چگالی احتمالی متمرکز شود به سمت تحلیل حساسیت فاصله مینکوفسکی همگرا می‌شود. همچنین مشابه شاخص‌های حساسیت پیشرو می‌توان شاخص‌های حساسیت معکوس را به دست آورد. سرانجام با تحلیل حساسیت یک طرح ذخیره‌سازی آب با ابعاد بالای خروجی‌های مدل پرداخته می‌شود.

**کلید واژه‌ها:** تحلیل حساسیت کلی، تحلیل حساسیت فاصله مینکوفسکی، تحلیل حساسیت منطقه‌ای، رده‌بندی خروجی.

### ۱. مقدمه

امروزه با توجه به توسعه‌ی علم آمار در تجارت و صنایع تولیدی از پیش‌بینی آماری به‌عنوان ابزاری قدرتمند برای بیان رخدادهایی که ممکن است در آینده اتفاق بیافتد، استفاده می‌شود. مدل‌سازی آماری ابزاری قدرتمند برای توسعه و آزمایش تئوری‌ها از طریق توضیحات علی، پیش‌بینی و توصیف می‌باشد. علت استفاده از مدل‌سازی آن است که پدیده‌های اطراف ما چه طبیعی و چه مصنوعی، توسط قوانینی اداره می‌شوند که بایستی آن‌ها را کشف کرده و به‌صورت مدل‌های ریاضی از آن‌ها به نفع خود استفاده کرد [۱]. به‌طور مشخص ممکن است در هنگام پیش‌بینی، پارامترها بر عملکرد مدل‌های آماری تأثیر بگذارند. برای مثال هنگامی که عدم قطعیت در پارامترها وجود داشته باشد، باعث ایجاد عدم قطعیت در پیش‌بینی مدل نیز می‌شود [۲]. به‌طور کلی آنچه که مدل‌سازی و هدف پژوهش را بسیار پیچیده می‌کند عدم قطعیت می‌باشد. عدم قطعیت روش علمی نیست، بلکه ماهیت آن است [۳]. برای این منظور از روش حساسیت کلی استفاده می‌شود. در طول دهه‌های گذشته روش‌های مختلفی از تحلیل حساسیت کلی ارائه شده است، یکی از این روش‌ها تحلیل حساسیت واریانس مبنا است [۴، ۵]. از

آنجا که واریانس فقط خلاصه‌ای از کل توزیع احتمال را ارائه می‌دهد، هنگام استفاده از واریانس به تنهایی برای نشان دادن عدم قطعیت، منجر به از دست رفتن اجتناب ناپذیر اطلاعات می‌شود. روش تحلیل چگالی مینا ( $DSA$ ) را می‌توان به عنوان رویکرد جایگزین مناسب برای این روش اشاره کرد [۶]. از این روش برای اندازه‌گیری تأثیر ورودی مدل بر اساس تابع چگالی احتمال توام ( $PDF$ ) توزیع خروجی استفاده می‌شود. ابتدا از این روش برای مدل‌هایی که دارای یک خروجی بودند استفاده می‌شد [۶]. به مرور زمان این روش برای مدل‌های دارای چندین خروجی گسترش یافت [۷]، تا روش تحلیل حساسیت چگالی مینا به یک فرم کلی‌تر برسد. روش تحلیل حساسیت فاصله مینکوفسکی ( $MD$ ) یک روش تعمیم یافته از روش تحلیل حساسیت چگالی مینا از مرتبه‌ی  $p$  می‌باشد [۸]، هنگامی که مرتبه‌ی  $p = 1$  باشد تحلیل حساسیت چگالی مینا به دست می‌آید.

علاوه بر این، تحلیل حساسیت منطقه‌ای ( $RSA$ ) روش دیگری از روش‌های تحلیل حساسیت کلی می‌باشد. این روش برای اندازه‌گیری تأثیر خروجی مدل بر اساس تابع چگالی احتمال توام ( $PDF$ ) توزیع ورودی استفاده می‌شود [۹]. بر خلاف  $MD$ ، ایده اصلی  $RSA$  این است که خروجی مدل را به چند رده جداگانه تقسیم کرده و سپس توزیع شرطی هر پارامتر ورودی بین رده‌ها را مقایسه کند.

به غیر از روش‌های  $GSA$  که در بالا ذکر شد، روش‌های دیگری نیز وجود دارد [۱۰-۱۲]، آیا می‌توان روابطی بین روش‌های مختلف به دست آورد؟ در این مقاله، هدف بررسی روابط روش‌های حساسیت  $GSA$  و هم‌ارزی مجانبی بین آن‌ها است. در واقع، در هنگام فراموشی جهت علی از ورودی به خروجی مدل، ورودی و خروجی صرفاً مجموعه متغیر تصادفی وابسته با توزیع توام می‌باشند. به همین امر می‌توان از معیارهای وابستگی بین ورودی و خروجی برای تحلیل حساسیت در پیش‌بینی‌ها استفاده کرد.

این مقاله به شرح زیر تنظیم شده است. بخش ۲ مروری کوتاه بر  $MD$  و  $RSA$  دارد. بخش ۳ ابتدا رابطه بین روش‌های  $MD$  و  $RSA$  را ارائه می‌دهد و اثبات هم‌ارزی بودن مجانبی آن‌ها را ارائه می‌دهد. در بخش ۴، به بررسی و مقایسه این دو روش در یک مدل هیدرولوژیکی پرداخته می‌شود و در بخش ۵ به بیان نتیجه‌گیری از مطالب بیان شده در این مقاله پرداخته می‌شود.

## ۲. مواد و روش‌ها

$Y = g(X)$  یک مدل می‌باشد که در آن  $X = \{X_1, X_2, \dots, X_d\}$ ، یک بردار  $d$  بعدی از پارامترهای ورودی تصادفی است که آن را با  $PDF$  توام  $f_X(x)$  نمایش می‌دهند.  $PDF$  حاشیه‌ای  $x_i$  را نیز با  $f_{X_i}(x_i)$  نشان می‌دهند. خروجی‌های مدل  $m$  بعدی را می‌توان به صورت  $Y = \{Y_1, \dots, Y_m\}$  نمایش داد.  $PDF$  توام خروجی مدل را با  $f_Y(y)$  نمایش داد.

### ۱-۲. معیار فاصله مینکوفسکی

روش معیار فاصله مینکوفسکی، برای اندازه‌گیری تأثیر ورودی‌های مدل بر اساس تابع چگالی احتمال توام ( $PDF$ ) از توزیع خروجی استفاده می‌شود [۸]. از این روش برای مدل‌های دارای چندین خروجی از مرتبه‌ی  $p$  استفاده می‌شود. وقتی که متغیر ورودی  $X$  در مقدار مشخص  $x_i$  ثابت باشد،  $PDF$  توام شرطی خروجی مدل را می‌توان به صورت  $f_{Y|X_i}(y)$  نشان داد؛ بنا بر این میانگین تأثیر در خروجی مدل برابر است:

$$E_{X_i}(\sup_{y \in Y} |f_Y(y) - f_{Y|X_i}(y)|) = \int_{X_i} (\sup_{y \in Y} |f_Y(y) - f_{Y|X_i=x}(y)|) f_{X_i}(x) dx \quad (1)$$

هنگامی که  $p = 1$  باشد معادله ۱ برابر با شاخص حساسیت فاصله مینکوفسکی را می‌توان به صورت زیر بیان کرد:

$$\delta_i = \frac{1}{p} E_{X_i}(s(X_i)) \quad (2)$$

## ۲-۲. روش تحلیل حساسیت منطقه‌ای

$$C = C_1, \dots, C_M \quad (3)$$
$$F_{X_i|C_k}(x_i) = Pr(X_i < x_i | Y \in C_k) \quad (r)$$
$$d_{C_k,i}^{CDF} = \int_{\mathcal{Y}_i} |F_{X_i}(x_i) - F_{X_i|C_k}(x_i)| dx_i \quad k = 1, \dots, M \quad (5)$$
$$S_{i,M}^{CDF} = d_{C_k,i}^{CDF} P_{C_k}$$
$$d_{C_k,i}^{PDF} = \int_{\mathbf{X}_i} |f_{X_i}(x_i) - f_{X_i|C_k}(x_i)| dx_i \quad (6)$$
$$S_{i,M}^{PDF} = \sum_{K=\setminus}^M d_{C_k,i}^{PDF} P_{C_K} \quad (\text{v})$$

پست الکترونیک: [info.biomath@du.ac.ir](mailto:info.biomath@du.ac.ir) تلفن تماس: ۰۲۳۳۵۲۲۰۱۱۸-۰۹۱۰۲۹۳۶۱۹۱ سایت: <https://biomath.du.ac.ir/>

### ۳. مقایسه دو رویکرد

#### ۱-۳. تحلیل حساسیت پیشرو و معکوس

آمارشناسان، یک دلیل منطقی برای تحلیل حساسیت کلی و برآورد داده‌ها ارائه داده‌اند.  $DSA$  را می‌توان مورد خاصی از این تحلیل حساسیت تعمیم یافته نام برد که در آن  $\mathbb{P}_Y$  یک ویژگی احتمالاتی برای خروجی مدل است مانند واریانس یا  $PDF$  [۱۴]. همین‌طور می‌توان به طور مشابه ویژگی احتمالاتی شرطی برای خروجی مدل را با  $\mathbb{P}_{Y|X_i}$  نمایش داد. بنا بر این فاصله بین  $\mathbb{P}_{Y|X_i}$  و  $\mathbb{P}_Y$  را می‌توان به صورت زیر اندازه‌گیری کرد:

$$\gamma_i(x_i) = \zeta(\mathbb{P}_Y, \mathbb{P}_{Y|X_i=x_i}) \quad (۸)$$

همان‌طور که در رابطه (۸) مشاهده می‌شود، نمادهای  $\zeta(.,.)$  و  $\gamma_i(x_i)$  به ترتیب بیانگر عملگر داخلی و آماره‌ی داخلی می‌باشد. شاخص حساسیت را می‌توان به صورت زیر تعریف کرد:

$$\xi_i = E[\gamma_i(X_i)] = E[\mathbb{P}_Y, \mathbb{P}_{Y|X_i}] \quad (۹)$$

$\xi_i$  شاخص حساسیت کلی  $X_i$  بر اساس  $\zeta(.,.)$  می‌نامند. با در نظر گرفتن انواع مختلف عملگرهای داخلی، شاخص حساسیت در معادله (۹)، می‌تواند با شاخص‌های مختلف حساسیت مانند شاخص حساسیت واریانس مبنای، شاخص حساسیت چگالی مبنای و .... مطابقت داشته باشد.

تفاوت اصلی بین  $RSA$  و منطق رایج در تحلیل حساسیت بر اساس معادله (۹) این است که  $RSA$  بر ویژگی‌های احتمالی ورودی مدل متمرکز است در حالی که منطق رایج بر ویژگی‌های احتمالی خروجی مدل متمرکز است؛ بنا بر این،  $RSA$  را می‌توان به عنوان تحلیل حساسیت معکوس (ورودی گرا) و منطق توام را می‌توان تحلیل حساسیت پیشرو (خروجی گرا) دانست؛ پس هنگام تحلیل حساسیت می‌توان از ویژگی‌های احتمالاتی هر یک از ورودی یا خروجی‌های مدل استفاده کرد. یک چارچوب کلی برای شاخص حساسیت معکوس (ورودی گرا) مشابه با رابطه (۹) را می‌توان به صورت زیر بیان کرد:

$$\eta_i = E[E[\zeta(\mathbb{P}_{X_i}, \mathbb{P}_{X_i|Y})]] \quad (۱۰)$$

اگر از  $PDF$  به عنوان ویژگی احتمالاتی و معیار فاصله  $(L^1 - \text{نرم})$  در رابطه (۶) نیز به عنوان عملگر داخلی استفاده شود، رابطه زیر به دست می‌آید:

$$\eta_i^{PDF} = E_Y \left[ \int_{X_i} |f_{X_i}(x_i) - f_{X_i|Y}(x_i)| dx_i \right] \quad (۱۱)$$

با باز کردن رابطه‌ی بالا در نهایت رابطه‌ی زیر به دست می‌آید:

$$\begin{aligned} \eta_i^{PDF} &= \int_Y \int_{X_i} |f_{X_i}(x_i) - f_{X_i|Y}(x_i)| dx_i f_Y(y) dy \\ &= \int_Y \int_{X_i} |f_{X_i}(x_i) f_Y(y) - f_{X_i,Y}(x_i, y)| dx_i dy \\ &= \int_{X_i} \int_Y |f_Y(y) - f_{Y|X_i}(y)| dy f_{X_i}(x_i) dx_i \end{aligned} \quad (۱۲)$$

همان‌طور که مشاهده می‌شود، رابطه‌های (۲) و (۱۱) با هم برابر هستند هنگامی که مرتبه  $p = ۱$  باشد. این برابری باعث ایجاد ارتباط کلی‌تر بین روش‌های تحلیل حساسیت معکوس و پیشرو بر اساس مفهوم منطق مشترک می‌شود [۱۴].

پست الکترونیک: [info.biomath@du.ac.ir](mailto:info.biomath@du.ac.ir) تلفن تماس: ۰۲۳۳۵۲۲۰۱۱۸-۰۹۱۰۲۹۳۶۱۹۱ سایت: <https://biomath.du.ac.ir/>

#### ۴. مورد آزمایشی: مدل آب‌شناسی

##### ۱-۴. شرح مدل

در این بخش به بررسی دو روش تحلیل پیشرو و معکوس در یک مدل آب‌شناسی های‌مود<sup>۱</sup> پرداخته می‌شود. در مطالعات آب‌شناسی، اغلب از مدل‌ها برای کمک به مدیریت منابع آب، پیش‌بینی آب‌شناسی، پیش‌بینی تغییرات آب و هوا و غیره استفاده می‌شود. رودخانه کارون از مرتفع‌ترین قله سلسله جبال زاگرس سرچشمه می‌گیرد و به خلیج فارس می‌ریزد. شاخه کارون متشکل از سه شعبه عمده کارون علیا، بازفت و خرسان مساحتی در حدود ۲۴ هزار کیلومتر مربع را در محل پل شالو زهکشی می‌کند. داده‌های مورد نیاز برای انجام این تحقیق شامل بارش و تبخیر می‌باشد که داده‌های استفاده شده از سایت هواشناسی (https://en.tutiempo.net/climate/۰۵-۲۰۱۴/ws-۴۰۷۹۸۰.html) در یک سال آبی جمع‌آوری شده است. در این طرح از مدل‌سازی برای پیش‌بینی آب و هوا و از روش‌های حساسیت کلی به عنوان ابزاری برای توسعه مدل‌های زیست محیطی استفاده می‌شود. در این خصوص مطالعات زیادی انجام شده است که در این مقاله مورد بررسی قرار می‌گیرد. در این طرح با استفاده از پودمان‌های<sup>۲</sup> هواشناسی به جمع‌آوری داده‌ها پرداخته که منجر به تشکیل مدل "های‌مد" می‌شود. مدل "های‌مد" یک مدل آب‌شناسی برای شبیه‌سازی رواناب بارندگی است که از یک پودمان برف، پودمان رطوبت خاک و پودمان مسیرگزینی جریان تشکیل شده است [۱۶]. پودمان برف از طریق روش درجه‌ی روز، میزان ذوب شدن برف را محاسبه می‌کند [۱۷]. هنگامی که متوسط دمای فصل پائیز کمتر از متوسط دما برای برف برسد ذخیره‌سازی برف اتفاق می‌افتد. هنگامی که متوسط دمای روزانه بالاتر از آستانه دما باشد ذوب برف اتفاق می‌افتد. پودمان رطوبت خاک، یک تابع توزیع معین برای ذخیره‌سازی حداکثر میزان رطوبت خاک ( $C_{max}$ ) برای عناصر حوضه آبریز می‌باشد. پارامتر نمایی  $\beta$  درجه‌ی خاصی از تغییرپذیری مخزن را نشان می‌دهد. رطوبت خاک در کسری بالقوه تبخیر می‌شود که با استفاده از روش "هامون"<sup>۳</sup> می‌توان آن را برآورد کرد. از باقی‌مانده بارندگی و ذوب برف برای پر کردن ذخایر رطوبت خاک استفاده می‌شود. مقدار اضافی بارندگی به پودمان مسیرگزینی جریان ارسال می‌شود که این پودمان میزان بارندگی اضافی را با ضریب  $\alpha$  افزاز کرده و از طریق مخازن خطی مفهومی موازی جهت شبیه‌سازی پاسخ‌های جریان سریع و آهسته سامانه را هدایت می‌کند. این ضریب سریع مخزن ( $R_q$ ) و آهسته مخزن ( $R_s$ ) است که جریان خروجی از هر مخزن را به جریان رودخانه کنترل می‌کنند. بنا بر این جریان شبیه‌سازی شده، مجموع خروجی‌های هر یک از این مخازن است [۱۸].

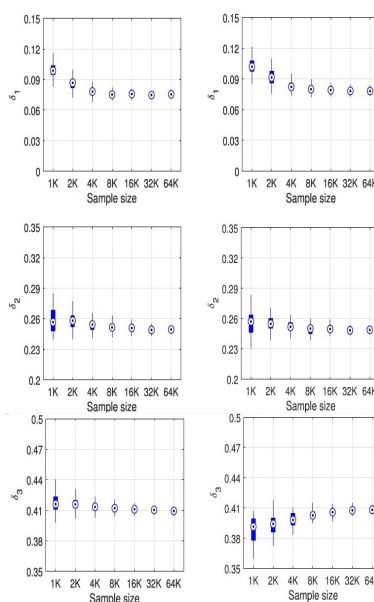
برای بررسی جامع بین دو روش شاخص حساسیت (شاخص معکوس و شاخص پیشرو) در مدل، ابتدا باید پارامترهای ورودی و خروجی مدل را تعیین کرد. در این مطالعه از اطلاعات بارش روزانه، تبخیر و تعرق به‌عنوان داده استفاده می‌شود. در این طرح سه پارامتری که روی مدل "های‌مد" تأثیر می‌گذارند عبارت‌اند از:  $\beta$ ،  $C_{max}$  و  $\alpha$  که می‌توان به‌عنوان پارامترهای ورودی آن‌ها را در نظر گرفت (جدول ۱)؛ بنا بر این یک بردار ۳ بعدی از پارامترهای ورودی تصادفی به‌صورت زیر به وجود می‌آید:

$$X = (X_1, X_2, X_3) \quad i = 1, 2, 3$$

جدول ۱: پارامترهای مدل "های مد"

پارامتر	حداقل	حداکثر
$C_{max}(X_1)$	۰	۲۹۷۰
$\beta(X_2)$	۰	۲
$\alpha(X_2)$	۰	۱

مقادیر و پارامترهای به دست آمده در جدول ۱ از وبسایت رسمی رود کارون به دست آمده است، همان‌طور که بیان شد طرح ذخیره‌سازی آب در یک دوره‌ی یک ساله مورد بررسی قرار می‌گیرد که مراحل بررسی آن به صورت زیر انجام می‌شود. برای مقایسه جامع دو روش پیشرو و معکوس با استفاده از جریان‌های شبیه‌سازی و با توجه به تعداد روزهایی که می‌خواهیم بررسی انجام شود، خروجی مدل از آن به دست می‌آید؛ بنا بر این با توجه به بررسی ۶۰ و ۱۰ روزه، خروجی مدل ۶ و ۳۶ بعدی است. برای تعیین تعداد رده مناسب برای روش‌های پیشرو و معکوس از  $N^{\frac{1}{2}}$  استفاده شود که در آن  $N$  اندازه کل نمونه است [۱۴]. در روش معکوس، نمونه‌های خروجی با استفاده از روش  $K - means$  به رده‌های مختلف تقسیم می‌شوند. همه این رده‌ها به صورت دوبه‌دو مجزا و اجتماع آن‌ها برابر با کل فضای خروجی مدل است. شکل‌های ۱ و ۲ برآورد شاخص‌های حساسیت را با توجه به افزایش اندازه نمونه‌ها برای خروجی‌های مدل با اندازه‌های به ترتیب ۶ و ۳۶ نشان می‌دهند.



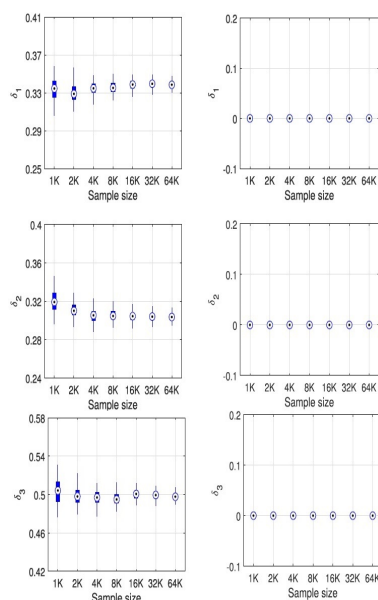
شکل ۱: نمودار جعبه‌ای برای خروجی‌های ۶ بعدی



جدول ۲: برآوردهای متوسط نهایی  $\delta_i$  (۶ پارامتری)

روش	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$
معکوس	۰/۰۷۱	۰/۲۴۴	۰/۴۰۴
پیشرو	۰/۰۷۳	۰/۲۴۵	۰/۴۰۵

نمودارهای سمت چپ در شکل‌های ۱ و ۲ بیانگر نتایج به دست آمده با استفاده از روش تحلیل معکوس می‌باشند [۱۹] و نمودارهای سمت راست در شکل‌های ۱ و ۲ بیانگر نتایج به دست آمده با استفاده از روش تحلیل پیشرو است [۱۴]. نمودارهای جعبه‌ای با ۳۰۰ بار اجرای مستقل از روش‌های تحلیل بیان شده به دست آمده‌اند. همان‌طور که مشاهده می‌شود، بالا و پائین این جعبه‌ها صدک‌های ۲۵ و ۷۵ مربوط به نتایج می‌باشد. شکل ۱ بیانگر این است که این دو روش حساسیت با افزایش اندازه نمونه برآورد مشابه‌ای را برای مدل‌های با تعداد خروجی ۶ نشان می‌دهند. همچنین جدول ۲ نتایج متوسط نهایی مربوط به این دو روش با اندازه نمونه ۲۰۰۰ را نشان می‌دهد. با توجه به نتایج به دست آمده در جدول ۲ می‌توان به این نتیجه دست یافت که هنگامی که مدل تعداد کمی خروجی داشته باشند این دو روش نتایج یکسانی را ارائه می‌دهند.



شکل ۲: نمودار جعبه‌ای برای خروجی‌های ۳۶ بعدی

در شکل ۲ نتایج مربوط به ۳۶ خروجی را نشان می‌دهد. روش معکوس همچنان با افزایش اندازه نمونه همگرا است؛ در حالی که در روش پیشرو برای همه‌ی پارامترها برآورد را نزدیک صفر نشان می‌دهد. لذا برای برآورد روش معکوس می‌بایست از  $PDF$ ‌های یک بعدی استفاده شود در حالی که روش پیشرو نیاز به



برآورد  $PDF$  های ۳۶ بعدی دارد. بنا بر این برآورد دقیق چنین  $PDF$  هایی با ابعاد بالا کار آسانی نیست و تفاضل میان  $PDF$  های شرطی و غیرشرطی را نمی‌توان به راحتی به دست آورد. از این رو برآوردهای روش پیشرو همواره نزدیک به صفر است.

## ۵. نتیجه‌گیری

با توجه به مطالب بیان شده می‌توان به این نتیجه رسید، هنگامی که تعداد خروجی‌های مدل زیاد باشد برای به دست آوردن نتایج می‌بایست از روش معکوس برای تحلیل مدل استفاده شود به این دلیل که در این روش از  $PDF$  های یک بعدی استفاده می‌شود در حالی که در روش پیشرو برای برآورد نیاز به  $PDF$  های چند بعدی دارد، همین امر باعث سادگی و بهتر بودن این روش نسبت به سایر روش‌های مشابه مانند فاصله مینکوفسکی، پیشرو، کپول تاک و ... است. همچنین لازم به ذکر است هنگامی که تعداد خروجی‌های مدل کوچک باشد لزومی ندارد از کدام روش استفاده شود به این دلیل که نتایج یکسانی را نشان می‌دهند.

## مراجع

- [1] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, f., Saisana, M., and Tarantola, S. (2008), Global sensitivity analysis: the primer, John Wiley Sons.
- [2] Watkins Jr, D. W. and McKinney, D. C. (1997), "Finding robust solutions to water resources problems," Journal of water resources planning and management, 123, 49–58.
- [3] Oreskes, N., Shrader-Frechette, K., and Belitz, K. (1994), "Verification, validation, and confirmation of numerical models in the earth sciences," Science, 263, 641–646.
- [4] Sobol, I. M. (1993), "Sensitivity analysis for non-linear mathematical models," Mathematical modelling and computational experiment, 1, 407–414.
- [5] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010), "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," Computer physics communications, 181, 259–270.
- [6] Borgonovo, E. (2007), "A new uncertainty importance measure," Reliability Engineering System Safety, 92, 771–784.
- [7] Lijie, C., Bo, R., and Ze, L. (2016), "Importance measures of basic variable under multiple failure modes and their solutions," in 2016 IEEE Chinese Guidance, Navigation and Control Conference (CGNCC), IEEE, pp. 1605–1611.
- [8] Zhai, Q., Yang, J., Xie, M., Zhao, Y. (2014). Generalized moment-independent importance measures based on Minkowski distance. European Journal of Operational Research, 239(2), 449-455.
- [9] Spear, R. C., Grieb, T. M., and Shang, N. (1994), "Parameter uncertainty and interaction in complex environmental models," Water Resources Research, 30, 3159–3169.
- [10] Greagar, G., Manohar, C. S. (2015). Global response sensitivity analysis using probability distance measures and generalization of Sobol's analysis. Probabilistic engineering mechanics, 41, 21-33.
- [11] Fort, J. C., Klein, T., Rachdi, N. (2016). New sensitivity analysis subordinated to a contrast. Communications in Statistics-Theory and Methods, 45(15), 4349-4364.
- [12] Xiao, S., Lu, Z., Wang, P. (2018). Multivariate global sensitivity analysis for dynamic models based on energy distance. Structural and multidisciplinary optimization, 57(1), 279-291.
- [13] Liu, Q. and Homma, T. (2010), "A new importance measure for sensitivity analysis," Journal of nuclear science and technology, 47, 53–61.
- [14] Borgonovo, E., Hazen, G. B., and Plischke, E. (2016), "A common rationale for global sensitivity measures and their estimation," Risk Analysis, 36, 1871–1895.
- [15] Da Veiga, S. (2015), "Global sensitivity analysis with dependence measures," Journal of Statistical Computation and Simulation, 85, 1283–1305.
- [16] Kollat, J., Reed, P., and Wagener, T. (2012), "When are multiobjective calibration trade-offs in hydrologic models meaningful?" Water Resources Research, 48.
- [17] Bergstroem, S. (1975), "The development of a snow routine for the HBV-2 model," Hydrology Research, 6, 73.
- [18] Wagener, T., Boyle, D. P., Lees, M. J., Wheeler, H. S., Gupta, H. V., and Sorooshian, S. (2001), "A framework for development and application of hydrological models," Hydrology and Earth System Sciences, 5, 13–26.
- [19] Xiao, S., Oladyshkin, S., and Nowak, W. (2020), "Forward-reverse switch between densitybased and regional sensitivity analysis," Applied Mathematical Modelling, 84, 377–392.

## Influence of Input Factors in Statistical Models on Response Probabilities of order $p$

1<sup>st</sup> Farzad Eskandari<sup>1</sup>, 2<sup>nd</sup> Mohammad Khorasani<sup>2</sup>

<sup>1</sup> Faculty of Statistics, Mathematical and Computer Department of Statistics, Allameh Tabataba'i University, Tehran, Iran

<sup>2</sup> Faculty of Statistics, Mathematical and Computer Department of Statistics, Allameh Tabataba'i University, Tehran, Iran

mokhorasani555@gmail.com

**Abstract**— In today's world, using the statistical modeling process, natural phenomena can be used to analyze and predict the events under study. In particular, when predicting, parameters affect the performance of statistical models. In many risk assessment issues, the presence of uncertainty in the parameters leads to uncertainty in predicting the model. Global sensitivity analysis is a tool used to show uncertainty and is used in decision making, risk assessment, model simplification and so on. Minkowski distance sensitivity analysis and regional sensitivity analysis are two broad methods that can work with a given sample set of model input-output pair. One significant difference between them is that minkowski distance sensitivity analysis analyzes output distributions conditional on input values (forward), while regional sensitivity analysis analyzes input distributions conditional on output values (reverse). In this dissertation, we study the relationship between these two approaches and show that regional sensitivity analysis (reverse), when focusing on probability density functions of input, converges towards minkowski distance sensitivity analysis (forward) as the number of classes for conditioning model outputs in the reverse method increases. Similar to the existing general form of forward sensitivity indices, we derive a general form of the reverse sensitivity indices and provide the corresponding reverse given-data method. Finally, the sensitivity analysis of a water storage design with high dimensions of the model outputs is performed.

**Keywords**— *Global sensitivity analysis, Minkowski distance sensitivity analysis, regional sensitivity analysis, Classification of output.*